CALIBRATION OF THE PRICE S SOFTWARE COST MODEL

THESIS

James C. Galonsky, B.S.
Captain, USAF

AFIT/GCA/LAS/95S-1

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GCA/LAS/95S-1

CALIBRATION OF THE PRICE S SOFTWARE COST MODEL

THESIS

James C. Galonsky, B.S.
Captain, USAF

AFIT/GCA/LAS/95S-1

Approved for public release; distribution unlimited

AFIT/GCA/LAS/95S-1

CALIBRATION OF THE PRICE S SOFTWARE COST MODEL

THESIS

Presented to the Faculty of the Graduate School of Logistics
and Acquisition Management of the Air Force Institute of Technology
Air University
In Partial Fulfillment of the
Requirements for the Degree of
Master of Science in Cost Analysis

James C. Galonsky, B.S.
Captain, USAF

September 1995

Acknowledgments

I would like to extend thanks to my thesis advisor, Professor Daniel V. Ferens, for suggesting and supporting a research topic that applies directly to my profession as a cost analyst. The experience I have gained from studying the PRICE S parametric model will benefit me greatly when constructing cost estimates for the Air Force. Prof. Ferens and the rest of the thesis committee were always available for assistance, and kept me motivated with their enthusiasm.

A special thanks goes to Jim Otte of Dayton PRICE Systems, who spent several afternoons providing me with individual instruction on the mechanics of PRICE S. Mr. Otte unselfishly donated his time in support of my research effort; this included accepting phone calls at home during non-business hours to answer questions. Without his help, this study would have been extremely difficult to complete.

Sherry Stukes of Management Consulting and Research, and Gina Novak-Ley of the Space and Missile Systems Center, deserve gratitude for providing me with a user friendly database for the calibration effort. Their suggestions and insight helped guide my efforts in a systematic, logical fashion.

Lastly, I would like to provide my sincere thanks to the United States Air Force for investing in my future by providing me the opportunity to conduct this research effort and study full-time at the Air Force Institute of Technology.

James C. Galonsky

Table of Contents

List of Figures

List of Tables

Table | Page

Abstract

As more Department of Defense resources are being allocated toward software development, the necessity to accurately plan for software costs has become critical. Obtaining reliable estimates from software cost models, like PRICE S, can be a problem when input parameters are not precisely defined or calibrated. This research effort centered on refining Productivity Factor (PROFAC) values for defense industry applications. The Space and Missile Systems Center Database was used to calibrate PROFAC values for eleven stratified data sets: military ground, military mobile, missile, unmanned space, Ada, Assembly, C, COBOL, FORTRAN, JOVIAL, and PASCAL. The accuracy of the calibrations was determined through comparisons of calibrated and default generated estimates versus actuals. Statistical methods used to make the comparisons included standard deviation, mean absolute error, mean relative error, and percentage of records estimated within twenty-five percent of actuals. The results were surprising in that, in most instances, the calibrated PROFAC values estimated actual cost well, but not overwhelmingly better than the default PROFAC values. The main contributing factor to this phenomena was variability within the stratified data sets. The results were encouraging, however, in that the results from seven of the eleven stratified data sets suggested either a new refinement in PROFAC values based upon the calibration or the recommendation to use PROFAC values from analogous calibrated records for estimating future efforts.

CALIBRATION OF THE PRICE S SOFTWARE COST MODEL

I. Introduction

General Issue

Software cost estimating has become an important current issue within the Department of Defense (DoD) due to the department's increased reliance on software to support weapon systems. Not long ago, cost estimating was directed mainly toward the hardware components of computer systems. Now that is changing as larger portions of system budgets are earmarked for the development and support of new software components. As more resources are being allocated toward software development, the necessity to accurately plan for software costs has become more critical (27:1-2).

A good way to accurately plan for software costs is to utilize a proven cost estimating model that yields reliable results. Cost analysts within DoD have an array of parametric cost models from which to choose: including PRICE S, REVIC, SASET, SEER, and SLIM. In the early stages of software development, these cost estimating models can provide the cost analyst with useful information before development costs are incurred (4:1).

Receiving useful information from cost models, however, can be a problem. Without taking the time to precisely define input parameters, the cost analyst will no doubt receive unreliable data and questionable cost estimates. Input parameters also can vary greatly from one model to the next. It is therefore appropriate for any cost analyst attempting to use a cost model to first be thoroughly familiar with the terminology, assumptions, estimating methodologies, strengths, and limitations of the model being used (4:3).

One popular model within DoD is the Programmed Review of Information for Costing and Evaluation- Software model (PRICE S). PRICE S is an empirical model developed with the combined experience and input of government and commercial software developers (13:2-1). An empirical model is a model which uses algorithms that have been verified consistently accurate through repeated observation and experimentation. The model was specifically created to assist project managers in assessing values for cost, time, and manpower based on the historical data of previous projects (9:1). PRICE S is commercially available from PRICE Systems, Moorestown NJ and the Air Force has an Air Force-wide license to use the software.

Specific Problem

It is crucial that a cost estimating model like PRICE S be properly calibrated to the users environment to obtain reliable cost data. Calibration is a process which adapts a general cost estimating model to a specific environment using historical information. It is not enough that a cost analyst be familiar with the functionality of a cost model. The analyst must also know the procedure to fine tune it properly to make the model as useful as possible. Very little effort has been expended in this pursuit within DoD. Because of this lack of proper calibration, many cost models that are used throughout DoD are providing analysts with results that are questionable.

In 1990, the Space and Missile Systems Center (SMC) initiated a contracted effort through Management Consulting and Research, Inc. (MCR) to calibrate the PRICE S, SEER, and SASET models to several environments: space, ground, avionics, and internally developed systems. SMC, a primary DoD software customer, benefited from the work done by MCR in improving the credibility of the cost models (13:1-1). But since 1991, the database that MCR used to calibrate the models has grown to 2616 data records

(14:2). With the increase in data records comes the opportunity to refine the model calibration efforts further.

Research Objective

The objective of this research effort is to refine the PRICE S calibration efforts to the military ground, missile, military mobile, and unmanned space environments; particularly through the discovery of applicable Productivity Factors (PROFAC) values for the various environmental platforms *and* software programming languages.

PROFAC Defined

The Productivity Factor (PROFAC) is a parameter which reflects the sum total of all individual talents, capabilities, and experience levels within an organization or industry. According to the developers of the PRICE S model, these attributes are among the most significant in determining the potential costs of developing software within any particular organization. Simply stated, a PROFAC value is a historically derived measure of relative performance (21:11-33). This measure of relative performance is a significant unknown cost driver of PRICE S for many organizations, and therefore, the most appropriate parameter to calibrate. Figure 1.1 shows the default values for PROFAC based on source language, application, and platform (17).

Productivity Factor, (PROFAC)

PROFAC 4.50

Source Language	Application	Platform	PROFAC
FORTRAN, COBOL, BASIC, 4th Generation, UNIX-C, PASCAL	2.0 - 4.0	0.8 - 1.0	9.0 - 6.0
FORTRAN, PASCAL, UNIX-C, BASIC, ATLAS	3.0 - 6.0	1.2	6.5 - 5.5
FORTRAN, ASSEMBLY, PL-1, MICROCODE, CMS-2, ALGOL	4.0 - 8.0	1.4	6.0 - 5.0
FORTRAN, PASCAL, ADA, ASSEMBLY, JOVIAL, CMS-2	3.0 - 9.0	1.8	5.5 - 4.5
FORTRAN, PASCAL, ADA, ASSEMBLY, JOVIAL	3.0 - 9.0	>1.8	5.0 - 4.0

FIGURE 1.1 Default PROFAC values

APPL Defined

The Application Mix (APPL) is a single parameter that takes into account the various amounts of each type of code in a particular set of software instructions. It is a single number that ranges form 0.866 to 10.952. Numbers toward the lower end of the range reflect software that is relatively simple to develop such as string manipulation and math operations. High end values reflect more complex software programs that are geared toward real-time and interactive applications. In general, APPL is a complexity gauge for the software code (21:11-33). The APPL value has a significant impact on determining how much effort is required to complete a given amount of source code; however, this value can range inconsistently from 0.866 to 10.952 for any given project; regardless of platform or language. It is therefore inappropriate, according to the model

developer, to presumptuously determine a calibrated APPL value to be applied to a particular environment. Each project must be individually evaluated in determining APPL values. Figure 1.2 shows the values of the APPL Generator used by PRICE S (17).

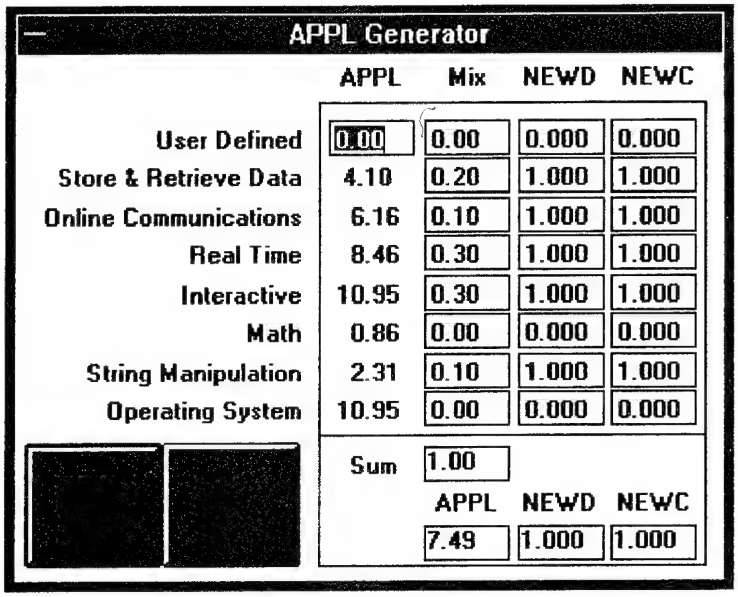

FIGURE 1.2 APPL Generator

Scope of Research

The research is being conducted to support SMC's goal of obtaining the latest calibrated PROFAC values for various environmental platforms and languages from the information contained in the SMC Database. The results will be considered for use in estimating the cost of new development efforts using the PRICE S parametric model. The study is restricted in that it will only evaluate records contained in the SMC Database. The research will be conducted as follows:

1. Learn how to use the PRICE S model and the SMC Database.

2. Stratify the SMC database into data sets suitable for PROFAC calibration.

3. Calibrate the PRICE S model to each of eleven stratified data sets: military ground, military mobile, missile, unmanned space, Ada, Assembly, C, COBOL, FORTRAN, JOVIAL, and PASCAL.

4. Determine the accuracy of the calibration experiment through comparison of calibrated PROFAC generated effort versus actual effort and default PROFAC generated effort versus actual effort for each suitable data record and data set.

5. Provide calibration findings and report conclusions.

This research effort differs from the prior effort by SMC in that more data points from recently completed software developments are available for calibration; hence, the opportunity to increase upon the accuracy of the prior study exists. The prior effort summarized comparisons of PROFAC calibration expectations to actuals for various platforms, but did not compare expectations to actuals for the various languages used in the database. The platform calibrations were simply broken down into languages for reference purposes. This study will compare calibration expectations to actuals for both environmental platform and language stratifications. The information will be presented using easy to understand parameters for accuracy such as Conte's "percentage within twenty-five percent" rule. This method involves counting how many calibrated and default PROFAC generated efforts within a given stratified data set fall within twenty-five percent of actual effort. This number is then divided by the total number of data points in the stratified set. The result is a comprehensive accuracy value for a given data set such as "accurate to within 25%, 65% of the time."

II. Literature Review

Overview

This chapter discusses background information regarding current views, limitations, and some pitfalls of software cost estimating. The chapter also provides a synopsis of two calibration efforts of the PRICE S software cost estimating model by 1) PRICE Systems and 2) MCR.

Software Cost Estimation

Barry Boehm, the author of Software Engineering Economics, points out that no software cost estimating tool or program is infallible. Having a parametric cost model like PRICE S does not guarantee correct software estimates. Boehm reinforces the fact that putting garbage into a cost model only results in receiving garbage out as a cost estimate (2:308).

Boehm suggests a six-step approach to software cost estimating that helps a cost analyst prevent putting invalid data into a cost model. His six steps are displayed in Figure 2.1 (2:309).

When establishing objectives, Boehm is mainly concerned that data gathering efforts are undertaken to support a specific decision that has been made. Wasted effort can result from undertaking a project without having clear objectives. If objectives remain unclear, the potential exists for data gathering requirements to change drastically and become costly (2:310). Objectives that are established must also be consistent with the needs of decision-makers who are going to use the cost estimating results. Those who are doing the data gathering must constantly balance the cost of obtaining information to the benefit of the information to the decision-maker without waste (2:312). In establishing objectives, it is also helpful to define the level of accuracy required in an estimate.

Either/or decisions may just require relative (less costly) estimates while other development decisions may require a single point estimate.

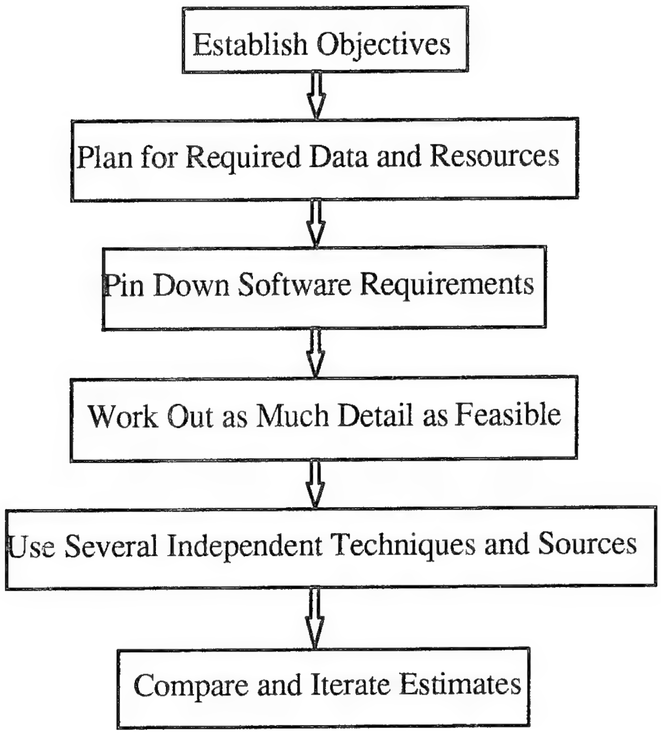

FIGURE 2.1 Six-Step Approach to Software Cost Estimating

Estimators must also plan for required data and resources. The time to perform an estimate generally is not the day before it is due. If an estimate is viewed as an important project, it becomes much more likely that the required personnel and time will be devoted to the project. More times than not, this will result in more accurate cost data. Boehm suggests that several things be kept in mind when planning for resources (2:314):

- when the estimate is due
- who is going to be responsible for each facet of the estimate
- how the project will be completed
- what tools will be used
- how much money is needed to complete the job

It is important to ensure that the estimator knows the exact software specifications of the project. Specifications should be testable, unambiguous, and quantitative. Extra effort may be warranted to ensure certain requirements are testable; but not all requirements will lend themselves to testability. One must always keep in mind the cost/benefit relationship when developing specifications. When specifications remain unclear, it may be wise to make conservative assumptions regarding that area of development rather than wasting too much time on the specification (2:316).

As much detail should be worked out during the estimation as possible. The more detail covered, the more everyone understands exactly what is going on. Also, the more detail that is covered in a work breakdown structure, the more likely it is that estimates will be generated from the lowest level possible in the WBS; leading to precision estimates that have lower variances than the norm. From an intuitive standpoint, the more detail entered into during estimation, the less likely it is to overlook important system costs to be included in the estimate (2:316). Details that remain unknown can be estimated using parametric methods.

Boehm suggests that analysts also utilize several independent techniques and sources in cost estimation. These techniques acts as good second opinions or "sanity checks" to an original estimate. Although a particular estimation technique may have been followed to the letter, the estimate can still fall prey to the inherent weaknesses in the particular technique chosen. Using several techniques to predict cost increases the probability that the various techniques will compensate for the others' weaknesses and allow the estimator to hone in on the closest value for cost (2:323).

Once several techniques have been utilized, it is smart to compare and iterate the estimates. The question must be asked as to why each technique yields a different estimate. Going through this process helps in discovering which aspects of cost each technique overlooks and/or treats differently. The process also helps in discovering cost

estimating error due to individual analysts (i.e. getting different results using the same technique). When estimate results are compared, particular attention must be given to components that make up the prime product; items that drive cost up the most. These are the components that can tolerate variability in estimates the least (2:324).

Lastly, Boehm suggests that frequent efforts to follow-up on data gathering should occur. The premise is that no data that has been used for analysis is flawless. The more follow-up data gathering that takes place, the more the gap can be closed between estimated and actual costs; given that the follow-up data represents improved information.

Frederick Brooks, author of the Mythical Manmonth, takes a humorous and somewhat pragmatic view to the cost estimating process as opposed to Boehm's more classical approach. According to Brooks, software projects are almost certain to exceed estimated schedule time. He contends that current techniques for software estimating are poorly developed; most of them take an overly optimistic approach to estimating. Secondly, he blames estimating techniques for confusing effort with progress (interchanging the terms man and month). Lastly, he believes that progress of development efforts are poorly monitored, and when any schedule slippage is discovered, a knee jerk reaction occurs which causes management to assign more and more manpower to a project; exacerbating the problem (3:14).

One of Brooks' key points is his contention that many times "men" and "months" are wrongfully used interchangeably in software estimation. He agrees that cost does fluctuate as a function of the number of personnel assigned and the number of months in project duration. His point, however, is that progress is not a simple function of man and month (3:16). In software development, many times a task cannot be divided effectively among workers due to the constraint of sequential tasks. Frequently, manpower is waiting for one element of the project to finish before the next group can start on the next element. In this case, adding more manpower to the project (effort) has little effect on the

schedule. Brooks likens the addition of manpower to solve schedule problems to assigning extra women to help a woman bear her child. Whether the woman remains by herself or has several women assigned to her, it will still take nine months to bear the child! (3:17)

Brooks also sees the difficulty in estimating due to personnel force structure; specifically the huge disparity of skill level from worker to worker. He cites a study by Sackman, Erikson, and Grant which measured performance levels of experienced programmers. The study found that ratios between best and worst performances averaged 10:1 and 5:1 on program speed and space measurements (3:30).

Given this information, one would assume that project managers would prefer to hire small, highly skilled teams as opposed to large, average skilled teams that get in each other's way when developing software. This ascertation is true; however, project managers run into a dilemma when hiring manpower. Brooks maintains that small, highly skilled teams are best for working on small projects, but the same is not true when a large project is undertaken. Many times a development effort is so large that a small, highly skilled team would take far too long to complete the effort. Sometimes the only option is to hire several workers with average ability for the project to make sure schedules are met (3:30, 31).

Obviously, a mixture of talent on a development effort makes estimation of cost more difficult. For a small project that is accomplished by a small, specialized team, variances in cost estimation tend to be much lower. As mentioned before, however, small projects are often the exception rather than the norm. Techniques must be found to estimate costs better when so much uncertainty exists due to worker skill level, project complexity, and desired schedule. This is why it is so important to have the aid of a software cost estimating tool such as PRICE S to help in making cost estimating decisions. Uncalibrated, PRICE S can handle a wide range of estimating for development

efforts with generally reasonable accuracy. But, as Barry Boehm would recommend, any technique used to estimate cost should be refined as much as possible to reduce variances from estimated to actual costs.

This is where the benefit of calibrating PRICE S to the user's particular environment is realized. One such calibration effort was undertaken in 1984 by PRICE Systems for the Deputy for Reconnaissance and Warfare, Aeronautical Systems Division (ASD/RW); now the Aeronautical System Center. The effort was initiated as part of the ASD/RW Cost Improvement Plan. The main objective was to increase acquisition management productivity and effectiveness (23:1).

The programs that were identified in the calibration effort were mainly aircraft avionics programs. PROFAC values were calculated for airborne military specification (mil spec) avionics, mil spec ground software, and production center (contracted) software programs. Calibration efforts yielded promising results as the standard deviation among calculated PROFAC values for specific platforms and languages was very small. For example, military ground software programs written in FORTRAN yielded a mean PROFAC of 3.81 with a standard deviation of 0.30 for thirteen data points (23:40). The small standard deviation between calculated PROFAC values increased confidence in the assumption that mean PROFAC values could be derived for specific platforms and successfully applied to future projects. This information was particularly useful since the results differed substantially from standard "uncalibrated" default values listed in the PRICE S user's manual. Generally, a military ground software program written in FORTRAN has a PROFAC value between 5.5 and 6.5 (21:C-6).

In 1990, Management Consulting and Research (MCR), with the guidance of PRICE Systems, embarked on a similar effort to aid the Space and Missile Systems Center (SMC) in calibrating PROFAC and APPL values from SMC's database. Four development environments were considered in the calibration:

- internally developed software (ID)
- military ground (MG)
- military air (MA)
- space programs (S)

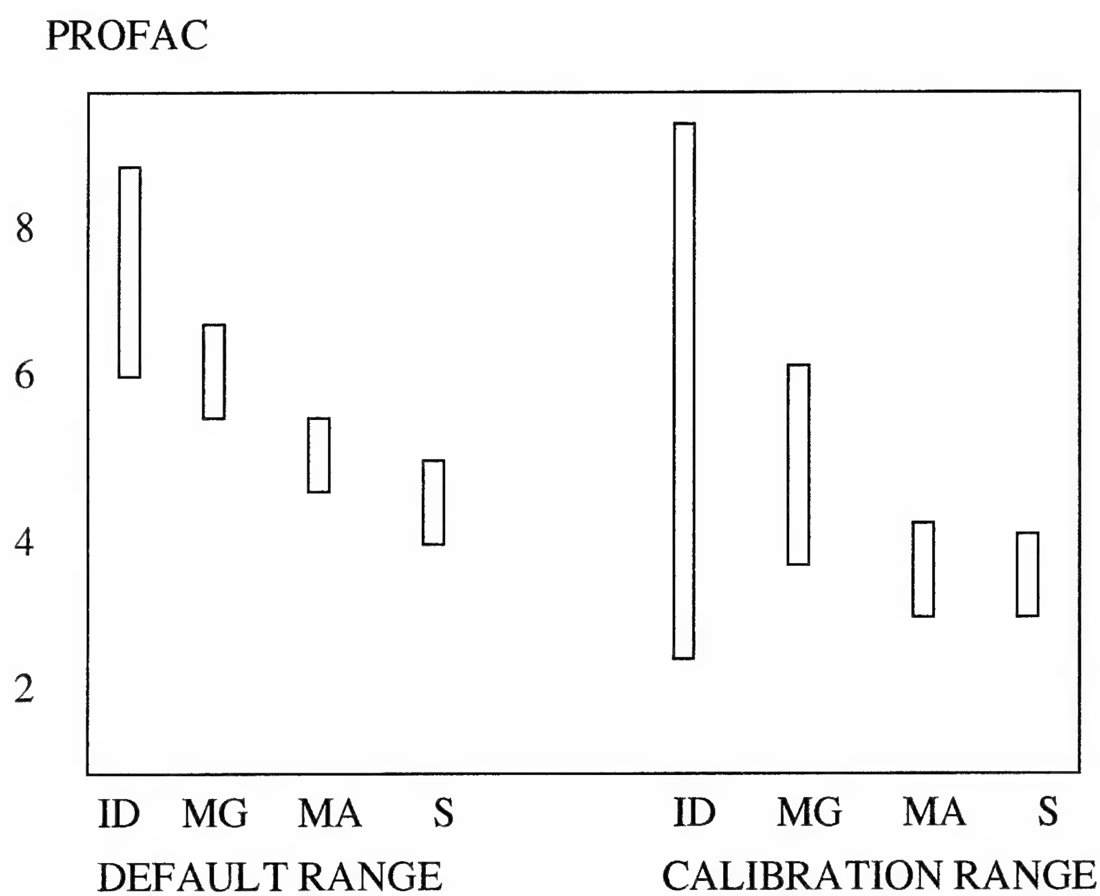

FIGURE 2.2 MCR Study PROFAC Calibration Summary

Results suggested that trends overlapped in the same range of PROFAC values initially published by PRICE Systems in their manual. The only exception was a wide deviation that was noted for internally developed software; attributable to the fact that there were only four data points (13:2-7). Figure 2.2 shows the range of PROFAC values observed for each environmental platform.

APPL values, as expected, were unpredictable; ranging from 0.87 to 10.95. Volatility in APPL values was particularly noted for the military ground environmental platform as shown by Figure 2.3.

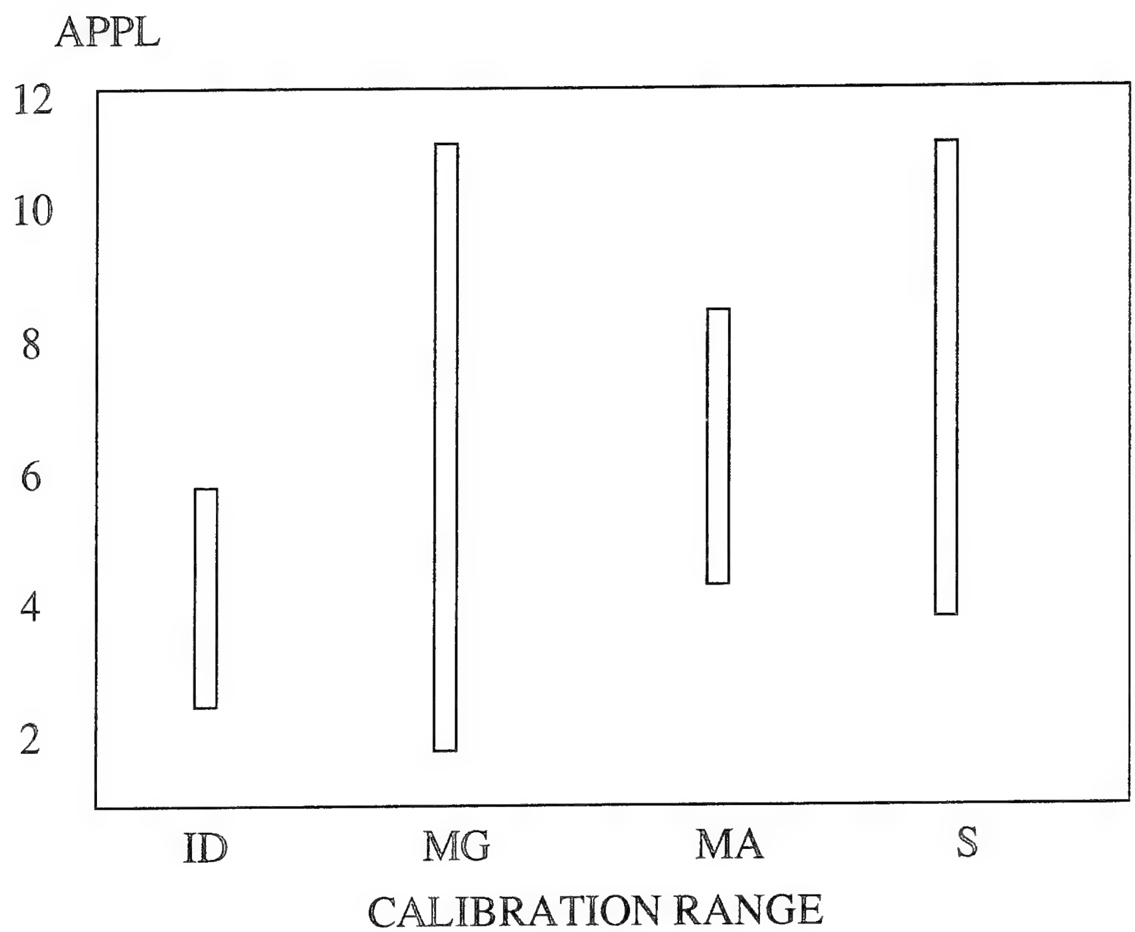

FIGURE 2.3 MCR Study APPL Calibration Summary

MCR's recommendation was to use calibrated PROFAC values rather than default values in the absence of specific information regarding an appropriate PROFAC for a new development effort. MCR further recommended that PRICE S calculated values for APPL based on code mix be used rather than the values obtained from the calibration study (13:2-9).

Summary

Barry Boehm contends that no software cost estimating tool or program is infallible; including PRICE S. This does not mean, however, that the outputs of each estimating method are unusable. Each independent technique has its own strengths and weaknesses and should be used in concert with other techniques as second opinions to original and consecutive estimates.

Although no estimating technique yields exact results, there is an opportunity to refine the accuracy of each methodology. The attempts by ASD/RW and MCR to calibrate the PRICE S model are two such examples of successful refinements.

III. Methodology

Overview

The methodology for calibrating PRICE S PROFAC values to the defense environment will be similar to the effort undertaken by MCR in 1990. Once again, the "organization" to be calibrated will not be a single organization, but the defense software industry as a whole. All records in the database will be combined when running PRICE S in the ECIRP calibration mode. The goal is to arrive at an industry PROFAC.

PROFAC values will, however, be separately calculated for different environmental platforms such as military ground, military mobile, missile, and unmanned space applications. Separate PROFAC values will also be determined for various higher order languages since the variability in software development costs can be great depending on which language is used (in PRICE S). The languages to be evaluated are Ada, Assembly, C, COBOL, FORTRAN, JOVIAL, and PASCAL.

All of the data needed for the calibration effort will be obtained from the SMC Database. Each of the 2616 records will be individually reviewed and investigated for use in the calibration effort. However, PRICE S was designed to respond to well-defined input parameters that may not be evident in every database record. One specific example would be the absence of cost or effort values for a record. In a case like this, the record is disqualified from the study. Other examples of potential problem areas are records which are unclear regarding schedule length or schedule phases included in the development effort. For all cases except the absence of values for effort, attempts will be made to make adjustments; with the assistance of the model developer, so the record can be useable for the study.

About the Database

The SMC Software Database uses an automated user interface to allow for access and display of data stored in the SMC Database. The database was developed under the direction of the USAF Space and Missile Systems Center directorate of cost (FMC), with assistance from the Space Systems Cost Analysis Group (SSCAG), and currently contains 2616 data records (14:2).

The database user interface is Windows-driven and allows for two methods of viewing records. One way is through browsing. The BROWSE function allows the user to begin at record 1 and page through all of the database records. An icon labeled "next" sends the user to the next record while an icon labeled "previous" sends the user back one record. When inside each record, the user also finds several arrow keys to help navigate the record. The arrow keys that contain a slash send the user to the beginning or the end of the record while the arrow keys without a slash bring the user forward or backward one page. The second method of viewing records involves using the FIND function to select a record based on record number and functional description. Once a record is selected, the user is automatically placed in the browse mode beginning with the record number that was selected (14:7,8).

A QUERY function allows the user to define criteria for a search and to receive a report based on that search. The user can define criteria such as software level, operating environment, application, software function, programming language, and confidence level. This has an advantage over the BROWSE and FIND functions in that it displays only the records that are of interest to the user; not the entire database (14:9).

Whichever function the user chooses, the result is the same once the individual record is viewed. Every record contains a five-section format (14:1-1 thru 5-2):

- Proprietary Data
- General Information
- Cost, Size, and Schedule Information
- Software Characteristics
- Maintenance Information

The records in the database and the information contained in each record are designed specifically to support input values required by several software estimating models; including PRICE S. The specific input parameters necessary for PRICE S can be carefully gleaned from the records. For instance, program globals can be found in Section 3 of each record and are used to customize the calibration to a specific schedule and labor profile (13:2-4).

Important data file "input variables" can be found throughout each record; mainly Pages 2, 3, and 4. These input variables can be translated to specific PRICE S input parameters such as effort, management complexity, internal integration, software language(s), complexity of the language(s), and APPL. This information can be used, in turn, to generate PROFAC values (calibrating the model).

Theory of Price S Calibration

Calibration is defined as an iterative process that adapts a model to new situations or new development efforts. The chief method of accomplishing this adaptation in PRICE S is by running the model in the reverse, or ECIRP mode (PRICE spelled backwards). This is performed by inputting actual cost, schedule, or effort data into the model and allowing the model to empirically solve for PROFAC values. The PROFAC value that results from running the model in the ECIRP mode is the first step in the calibration process (8:1).

The next step is to run the model in the forward mode, using the intial (ECIRP) PROFAC estimate. The cost estimate that is obtained from this forward run is then compared to the actual cost data that was entered as an input during the ECIRP calculations. If the two estimates are close, within a specified tolerance level (0.05%), the new PROFAC value is accepted and the calibration flow stops (8:1).

If the comparison yields a result that is unacceptable, the result is compared to previous results to determine if successive iterations are actually converging on more accurate PROFAC values. If the values are converging, the iteration process continues until resultant cost estimates are within tolerance. If the values are not converging

properly, the process is halted, and the calibration effort is deemed unsuccessful (8:1). The automatic ECIRP calibration process, as defined by PRICE Systems, is shown in Figure 3.1.

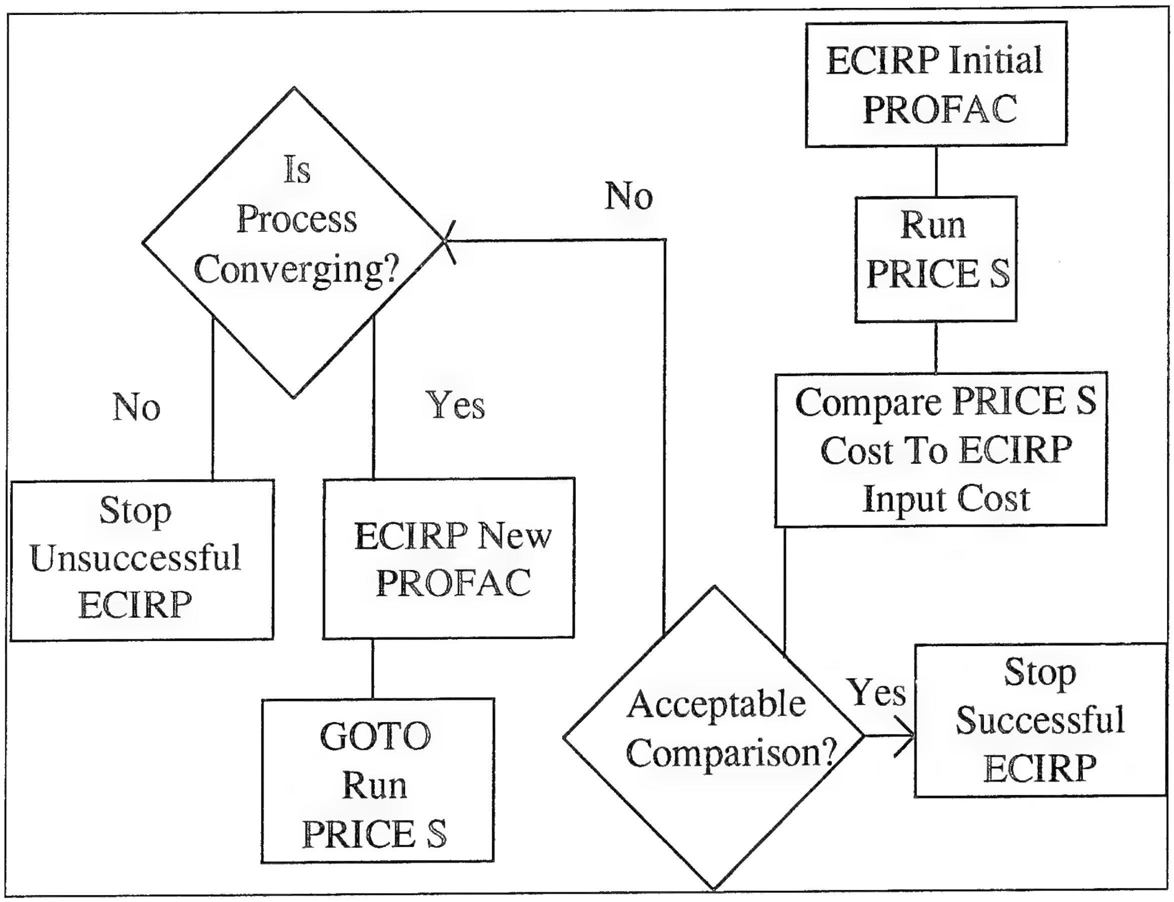

FIGURE 3.1 PROFAC Calibration Process

When calibration is applied to a number of software programs, however, a pattern will usually emerge, and improved PROFAC values can be obtained. The usefulness of refining PROFAC values lies in the fact that once a PROFAC is determined for an organization or specific environment, the value remains constant for quite some time (20:2E-1).

Step-by-Step PRICE S Calibration

After learning how to use the SMC Database user interface, the next step in the calibration effort is gathering information from the database to be entered into the PRICE S model. Since the database has record information which applies generally to several parametric cost-estimating models, it is useful to develop a data gathering form specifically geared to PRICE S (Fig. 3.2). Attributes that are not applicable to the SMC Database are shaded with X's.

1.0 Record Number____________________________
2.0 Application Name___________________________
2.6 Operating Environment________________________
2.7 Application_______________________________
3.1.1 Total effort (Person Months)___________________
3.1.3 Average Staffing Level______________________
3.1.4 Peak Staffing______________________________
3.2.1 New Unique SLOC__________________________
3.2.2 Common SLOC_____________________________
3.2.3 Reused SLOC______________________________
3.5.3 Schedule Completion Dates____________________

SRR	TRR
SDR	FCA
SSR	PCA
PDR	FQT
CDR	OTE
PQT (PRELIM QUAL)	

4.1 Level of Complexity____________________
4.2 Programming Languages__________________
Software Attributes
4.3.1 Application Complexity (APPL)___________
4.3.5 Display Requirements (APPL)____________
4.3.8 Memory Constraints (%)________________
4.3.9 Timing Constraints (%)_________________

FIGURE 3.2 PRICE S Data Collection Form

Attribute	V.Low	**Low**	**Nom**	**High**	**V. High**	**E. High**	**CPLX adj.**
4.3.3 Requirement Volatility	xxxxxx						
4.3.4 Rehosting Req.	xxxxxx	xxx					
4.3.6 Reusability Req.	xxxxxx	xxx					
4.3.7 Security Level	xxxxxx	xxx					
4.8.1 Personnel Exp.						xxxxxxxx	
4.8.2 Personnel Capabilities						xxxxxxxx	
4.8.5 Team Programming Exp.							
4.8.6 Dev. Methods Exp.							
4.8.7 Dev. Sys. Exp.							
4.8.8 Target Sys. Exp.							
4.23.4 Dev. Sys. Volatility	xxxxxx						
4.23.6 Quality Assur. Level						xxxxxxxx	
4.23.7 Test Level						xxxxxxxx	
4.23.8 Multiple Site Dev.	xxxxxx	xxx					
4.23.13 Modern Prac. Exp						xxxxxxxx	
4.23.14 Automated Tool Spt.						xxxxxxxx	
Net CPLX Effect							

Source Code Mix	**% of Total**	**% New Design**	**% New Code**	**% Retest**
Operating Sys				
Interactive Ops				
Real-Time C&C				
On-Line Comm				
Data S&R				
String Man.				
Math Oper.				

PROFAC Calibration result = ___________________

FIGURE 3.2.1 PRICE S Data Collection Form (Continued)

This handy form contains all the information needed to construct a record calibration in PRICE S. Once the entire database is searched for useable records and the information recorded for each record, the data collection forms can be filed into homogeneous data sets. This process is known as database stratification. Once data sets are established, the next step is to construct calibration files using the software model. The following is an explanation of the specific information that is entered into the model (using record 75 of the database as an example):

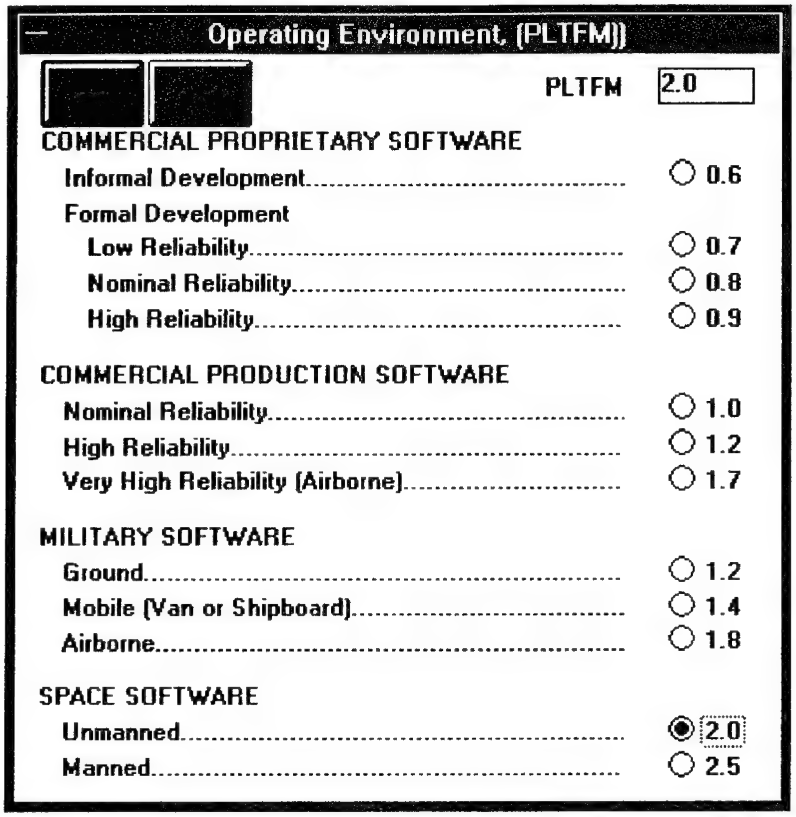

FIGURE 3.3 Operating Environment (PLTFM)

Operating Environment (Fig. 3.3) is a variable that describes the planned operating environment for a new development effort. It measures portability, reliability, structuring, testing, and documentation required by the customer. The most important issue in determining PLTFM value is the level of specification and testing involved in the effort (17). In this example, the PLTFM value is equal to 2.0 since record 75 contains information on unmanned space software.

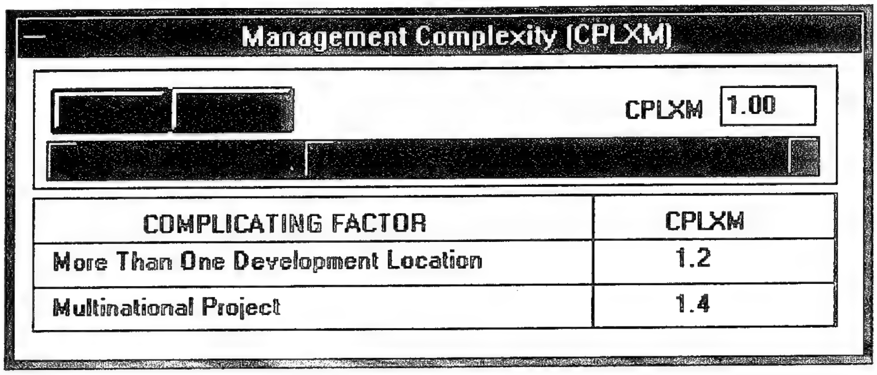

FIGURE 3.4 Management Complexity (CPLXM)

Management Complexity (Fig. 3.4) measures the relative effect of complicating factors on the entire software development effort. Values less than one have no effect on effort (17). In this example, CPLXM is equal to the industry average of one.

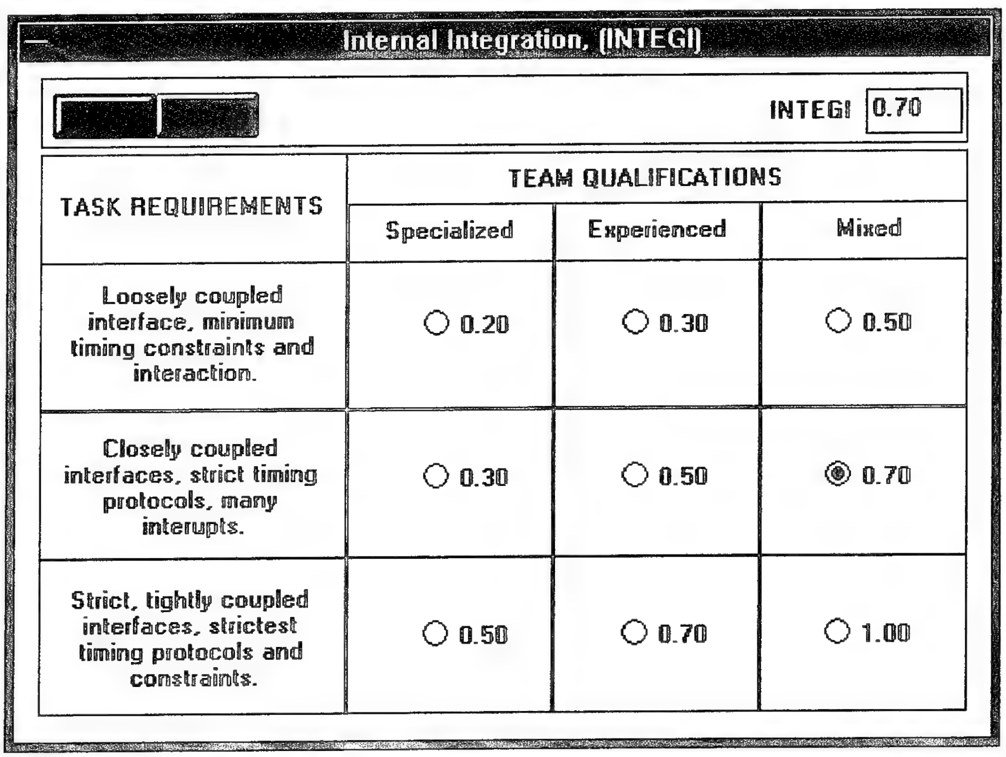

FIGURE 3.5 Internal Integration (INTEGI)

Internal Integration (Fig. 3.5) evaluates the level of integration and testing of components at the computer software configuration item (CSCI) level. 0.5 is the default value; numbers less than that reflect task requirements that are easier than average, and numbers greater than 0.5 describe more complex tasks (17). Team qualifications also have an impact on the INTEGI value. For example, average tasks performed by specialized teams warrant a lower INTEGI than the default value. Likewise, teams with mixed experience must add to the INTEGI value that would normally be selected based solely on task requirements. In this example, the INTEGI has a value of 0.70 to reflect a more complex integration task than average. This is mainly due to the "very high" requirements volatility listed in the record.

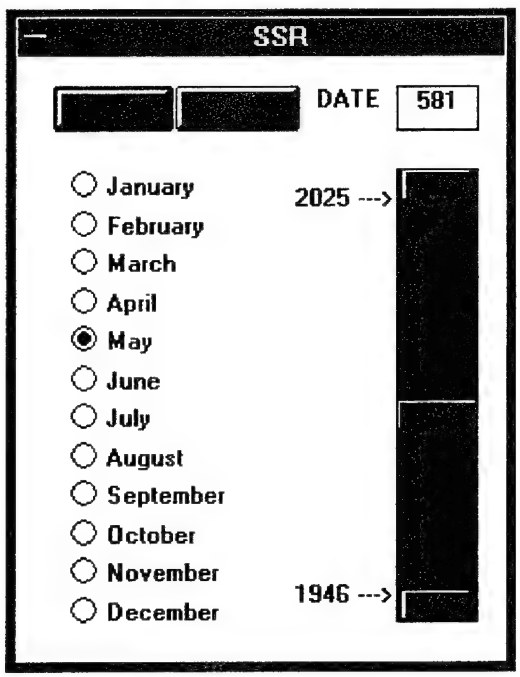

FIGURE 3.6 Software Specification Review

Software Specification Review (SSR) (Fig. 3.6) is the input parameter that specifies the completion date of the following tasks (17):

- interface requirements specification
- software requirements specification
- allocated baseline

Inputting either the SSR or Software Design Review (SDR) dates are mandatory because those values are used in the model to generate an internal reference schedule. The internal reference schedule constructed by the model acts as a basis for assessing penalties to estimated effort. Other schedule dates such as Functional Configuration Audit are helpful in computing internal schedules to be compared with the reference schedule in assessing penalties but are not mandatory. Dates which are unknown can simply be entered as a zero, but the model will still calculate values for those dates; consistent with either the SSR or SDR date that was entered into the model (21:11-4, 5). In this example, the SSR date is May 1981.

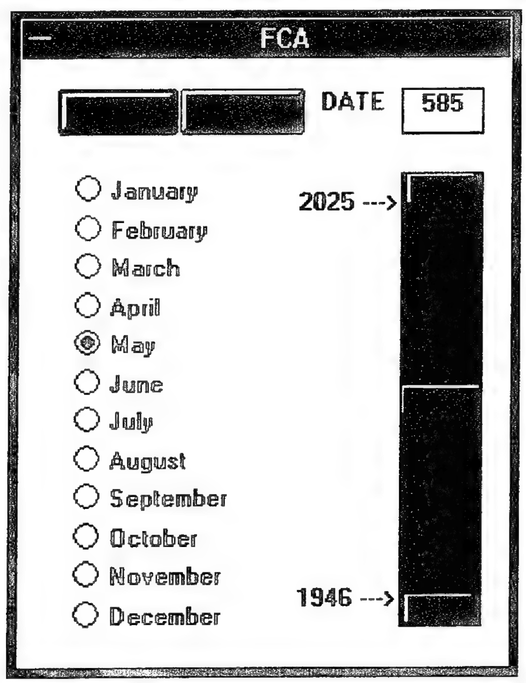

FIGURE 3.7 Functional Configuration Audit

The Functional Configuration Audit (Fig. 3.7) date is an input parameter highly suggested for inclusion in the model by the software developer. As mentioned before, it is a parameter that assists the model in calculating internal schedules to be compared with the reference schedule for possible effort penalty assessment. In this example, the FCA date is May 1985.

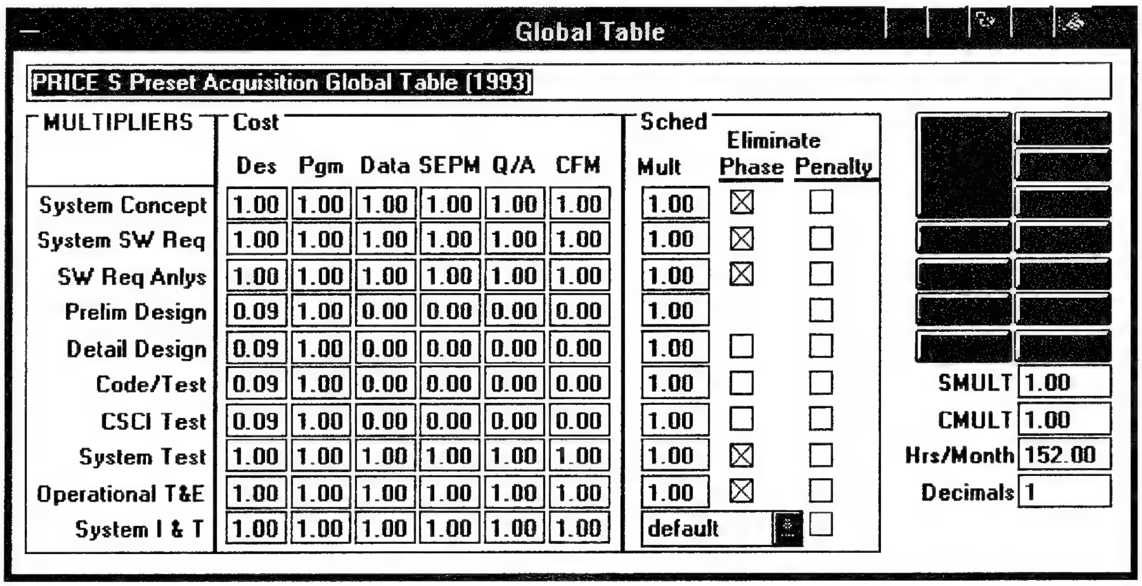

FIGURE 3.8 Global Table

The Global Table (Fig. 3.8) is comprised of four input sections: acquisition cost multipliers, acquisition schedule multipliers, phase eliminate indicator, and the penalty cost eliminate indicator. The acquisition cost multipliers are direct linear multipliers of all 54 cost elements depicted in the table grid by schedule phase and type of effort. The default values are set to one and can be increased or decreased based upon relative anticipated effort. Alternatively, an entire row (schedule phase) can be adjusted by changing the values in the acquisition schedule multiplier column (17).

The phase eliminate indicator allows for a schedule phase to be completely removed from calculations. Using the phase eliminate indicator is equivalent to placing a zero in the schedule multiplier column for that phase. And the penalty cost eliminate indicator allows no cost penalties to be applied to a specific phase. This negates any penalties that would normally be assessed through reference schedule calculations based on SSR or SDR dates (17). In this example, the effort applies to Preliminary Design through CSCI Testing; therefore, the five phases which do not apply to the record are eliminated from further cost calculations.

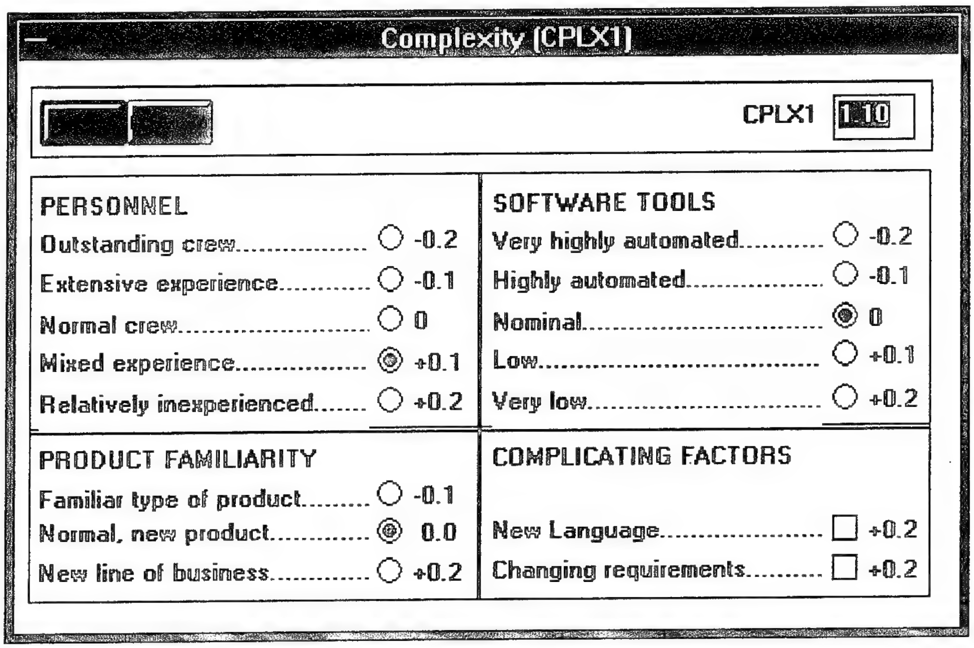

FIGURE 3.9 Complexity (CPLX1)

CPLX1 (Fig. 3.9) is similar to the Management Complexity parameter that was discussed earlier with the exception that it applies specifically to the complicating factors on the software development task itself. Personnel skills, product familiarity, software tools, and unusual complicating factors all affect development schedule and hence the CPLX1 value. CPLX1 is a performance variable that works in concert with the

acquisition mode parameters to aid in the calculation of a reference schedule. The parameter has a default value of one and can be increased or decreased based on the four factors pictured in the figure (17). In this example, the CPLX1 value is 1.1. "Low" team programming experience identified in the record penalizes (raises) the value by 0.1.

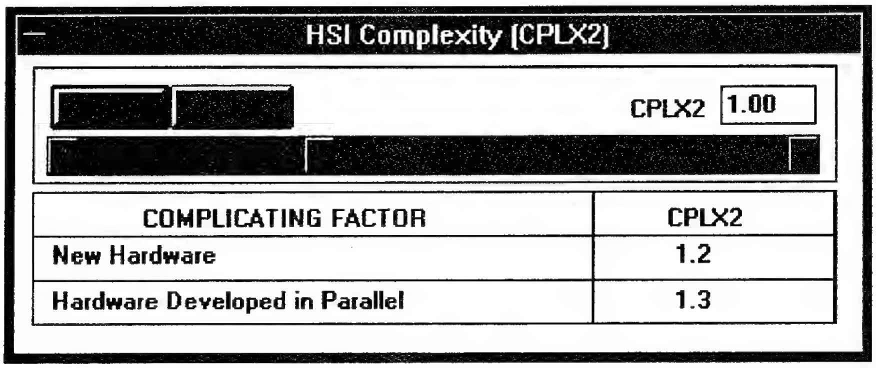

FIGURE 3.10 HSI Complexity (CPLX2)

CPLX2 (Fig. 3.10) is similar to CPLX1 in that it measures the relative effect of complicating factors on the software development task. It differs in that it is solely concerned with complicating factors caused by hardware/software interactions. The default value is one. New hardware used in the development task raises the value by 0.2; hardware developed in parallel with the software task raises the value by 0.3 (17). In this example, the CPLX2 value is the industry average of 1.0.

The APPL generator (Fig. 3.11) uses source code mix and percentage of new design and new code in determining an APPL value; a mandatory input. In this example, the source code mix yields a value of 7.49 for APPL.

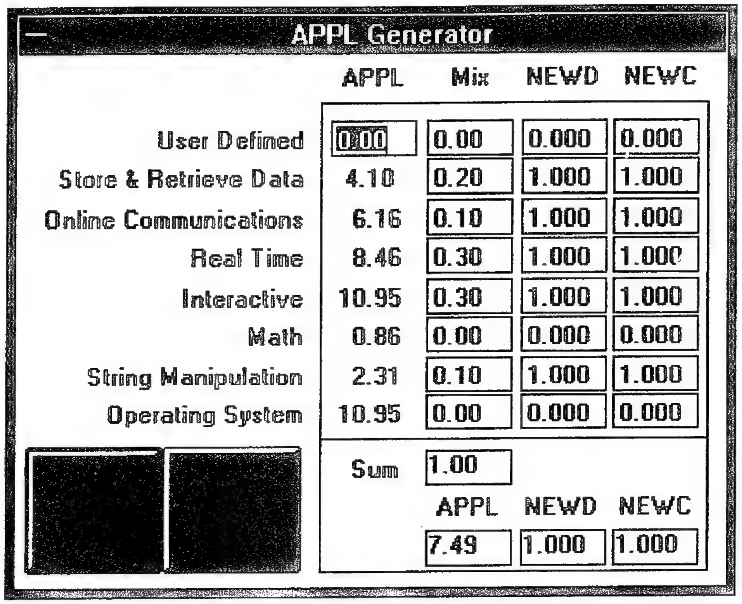

FIGURE 3.11 APPL Generator

The previous nine items, when combined with easy to understand input parameters such as language type, number of source lines of code, and person-months of effort, allow the user to calculate a PROFAC value for the record in the ECIRP mode automatically as previously described in the Theory of PRICE S Calibration found in Chapter III. The PROFAC value yielded by the example record #75 is 5.56. The actual person-months of effort for the record is 912. Normally, actual person-months of effort is an unknown parameter for which the analyst is trying to estimate. Running PRICE S in the normal forward mode using a PROFAC value of 5.56 yields a total effort of exactly 912 for the record; the actual effort. An exact calibrated PROFAC value, however, is not available for project estimates because the effort or cost is an unknown parameter at the beginning of a project. PROFACs can only be known with certainty once a project is completed and actual effort is known. In the absence of having an exact PROFAC value to use for a particular project and for a given situation, a default PROFAC value or previously calibrated PROFAC value from similar projects must be used to estimate effort. The

purpose of this research is to discover calibrated values for PROFAC that are more accurate than the default values when applied to projects of various operating environments and languages.

The next step in the calibration process is finding a calibrated PROFAC value for each record in the particular stratified data set. Once that is accomplished, the data set must be stratified further into a calibration set and a validation set. A validation set is simply a set of data omitted from the calibration set calculations that will be used in verifying calibrated PROFAC values. The rules for determining points for calibration and validation, as requested by SMC, the sponsor of this research, are shown in Table 1.

TABLE 1

DETERMINING POINTS FOR CALIBRATION AND VALIDATION

Number of Data Points	Points used for Calibration	Points used for Validation
8	ALL	NONE
9-11	8	THE REMAINDER
>=12	THE REMAINDER	ONE-THIRD

A random a sample is used in determining which points in a data set to use for validation. Once the data set has been segregated into calibration and validation points, the next step is to find the mean PROFAC value and sample standard deviation for all of the records in the set labeled for calibration. That mean PROFAC value becomes the testable calibrated PROFAC value for the entire platform or language data set. In instances where less than eight data points are available, testable calibrated PROFAC values will be calculated for **n** successive iterations, and one data point will be held for

validation in each iteration. The multiple iterations are an attempt to smooth out variability in results caused by having too few data points.

The final step in the methodology is to determine the accuracy of the uncalibrated model versus the calibrated model. This is done by running all of the calibration and validation records in the forward mode using both the calibrated PROFAC value and the uncalibrated default values. Effort values obtained from these forward runs are compared to actual effort and model accuracy is determined through the use of statistics including mean absolute error, mean relative error, and Conte's percentage within twenty-five percent rule described in Chapter I. The methods for computing mean absolute error and mean relative error are shown in Figure 3.12.

1. Mean Absolute Error = the Summation of (the Absolute Value of Calibrated or Default PROFAC estimated Effort minus Actual Effort) divided by the Number of Records in the Data Set

$$\frac{\left(\sum_{i=1}^{N} \left| EE_i - AE_i \right|\right)}{N}$$

2. Mean Relative Error = the Summation of ((the Absolute Value of Calibrated or Default PROFAC estimated Effort minus Actual Effort) divided by Actual Effort) divided by the Number of Records in the Data Set

$$\frac{\left(\dfrac{\sum_{i=1}^{N} \left| EE_i - AE_i \right|}{AE_i}\right)}{N}$$

EE= estimated effort
AE= actual effort
N= number of records in the data set

FIGURE 3.12 Mean Absolute Error and Mean Relative Error Computations

The distributions of the calibrated and default PROFAC estimated efforts are also compared to the distributions of the actuals and examined for biasedness using the Wilcoxon Signed-Rank Test. This non-parametric test checks for sizable differences between the sums of negative and positive deviations between estimated and actual effort in determining whether to reject or fail to reject the null hypothesis that no bias exists between the distributions. This is accomplished by comparing minimum deviations to T-statistics chosen by significance level. If the minimum deviations are greater than the T-statistics, the test fails to reject the hypothesis that no bias is introduced by the estimating distributions (15:680).

Summary

The methodology discussed throughout this chapter can be summarized in eight steps:

- Search SMC Database for records suitable for PRICE S calibration
- Collect information pertinent to PRICE S model inputs for each record
- Stratify records by environmental platform and language
- Segregate records in each stratified data set into calibration and validation points
- Discover individual calibrated PROFACs for each record using the PRICE S ECIRP mode
- Find the mean PROFAC of calibration points for each data set
- Estimate effort for each data set, by record, using mean PROFAC and default PROFAC
- Examine the results of the estimates, as they compare to actual effort, through various statistical methods

The next chapter summarizes the calibration results obtained through this methodology; presented by operating environment, language, and iteration in easy-to-read tabular format.

IV. Analysis and Findings

Overview

Seventy records from the SMC Database were found to be suitable for the PRICE S calibration effort. The input parameters for each record, organized according to operating environment and language, can be found in Appendix A. Default PROFAC values (Table 2) used in the study were determined by analyzing ranges published by PRICE Systems and through personal interviews with representatives from PRICE Systems. Only default PROFAC values applicable to the study are listed.

TABLE 2

DEFAULT PROFAC VALUES

	Unman. Space	Missile	Mil. Mobile	Mil. Ground	Avionics	Commercial
Platform Cals	4.5	5.0	5.5	6.0		
Ada	5.0	5.5	6.0	6.5	5.5	
Assembly	4.0	4.5	5.0	5.5		
C	5.0			6.5		6.5
COBOL				6.5		
FORTRAN	5.0		6.0	6.5		
JOVIAL	4.0	4.5			4.5	
PASCAL			6.0	6.5		6.5

Calibration Results

The calibration results for each operating environment, language, and iteration are presented in an easy to read format which includes the following:

- Records used in the calibration
- Mean PROFAC of calibration records
- Standard deviation of the PROFACs used in the calibration
- Mean absolute error for both mean and default estimated effort
- Mean relative error for both mean and default estimated effort
- Columns and rows that show record number, calibrated PROFAC for each record, actual effort as reported by the SMC Database, mean and default PROFACs of calibration records, estimated effort for each record using both mean and default PROFAC values, absolute deviation of estimated effort from actual effort, and absolute relative (percentage) error of estimated effort
- Accuracy within 25% of actual effort
- Narrative explaining the dynamics of each calibration result

In addition to the information presented for each calibration run, the Wilcoxon Signed-Rank Test was performed for all data sets containing eight or more data points. The results supported (failed to reject) the null hypothesis of no bias introduced by the distributions of estimated effort for every data set at the 0.10 and 0.05 levels of significance (see Appendix B).

TABLE 3

MILITARY GROUND

Military Ground

Records Used for Calibration: 2, 7, 9, 23, 26, 30, 50, 54, 56, 58, 63, 68, 73, 301, 2497, 2519, 2520, 2522, 2523, 2524

Mean PROFAC of Calibration Records:	6.12
Standard Deviation:	1.83
Mean Absolute Error:	
Of Effort using Mean PROFAC	329.32
Of Effort using Default PROFAC	327.41
Mean Relative Error:	
Of Effort using Mean PROFAC	29.41%
Of Effort using Default PROFAC	30.33%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
2	3.25	245.0	6.12	125.10	119.90	48.94%	6.00	128.10	116.90	47.71%
7	6.74	127.0	6.12	140.00	13.00	10.24%	6.00	142.90	15.90	12.52%
9	6.73	545.0	6.12	600.30	55.30	10.15%	6.00	613.20	68.20	12.51%
23	6.28	42.0	6.12	43.00	1.00	2.38%	6.00	43.90	1.90	4.52%
24	4.98	238.0	6.12	195.10	42.90	18.03%	6.00	198.70	39.30	16.51%
26	5.79	22.0	6.12	20.70	1.30	5.91%	6.00	21.10	0.90	4.09%
28	3.40	887.0	6.12	479.00	408.00	46.00%	6.00	488.70	398.30	44.90%
30	4.10	160.0	6.12	103.60	56.40	35.25%	6.00	105.80	54.20	33.88%
48	3.50	344.0	6.12	178.40	165.60	48.14%	6.00	182.50	161.50	46.95%
50	6.35	656.0	6.12	682.50	26.50	4.04%	6.00	697.20	41.20	6.28%
53	7.51	51.0	6.12	62.50	11.50	22.55%	6.00	63.70	12.70	24.90%
54	9.74	134.0	6.12	235.90	101.90	76.04%	6.00	241.40	107.40	80.15%
55	2.27	6713.0	6.12	2390.80	4322.20	64.39%	6.00	2439.80	4273.20	63.66%
56	7.34	23.0	6.12	27.90	4.90	21.30%	6.00	28.50	5.50	23.91%
58	4.12	639.0	6.12	419.70	219.30	34.32%	6.00	428.40	210.60	32.96%
59	6.07	147.0	6.12	145.70	1.30	0.88%	6.00	148.70	1.70	1.16%
61	9.31	113.0	6.12	186.90	73.90	65.40%	6.00	191.30	78.30	69.29%
63	7.19	23.0	6.12	28.00	5.00	21.74%	6.00	28.70	5.70	24.78%
68	6.94	170.0	6.12	197.20	27.20	16.00%	6.00	201.80	31.80	18.71%
72	8.54	42.0	6.12	60.50	18.50	44.05%	6.00	61.70	19.70	46.90%
73	8.09	56.0	6.12	75.20	19.20	34.29%	6.00	76.80	20.80	37.14%
301	4.88	430.0	6.12	339.30	90.70	21.09%	6.00	346.10	83.90	19.51%
2497	7.96	80.0	6.12	103.30	23.30	29.13%	6.00	105.10	25.10	31.38%
2501	6.78	418.0	6.12	467.70	49.70	11.89%	6.00	477.40	59.40	14.21%
2510	8.42	181.2	6.12	247.30	66.10	36.48%	6.00	251.90	70.70	39.02%
2519	8.22	1833.0	6.12	2547.90	714.90	39.00%	6.00	2602.30	769.30	41.97%
2520	2.44	3960.0	6.12	1575.70	2384.30	60.21%	6.00	1611.30	2348.70	59.31%
2521	3.27	735.0	6.12	388.90	346.10	47.09%	6.00	397.90	337.10	45.86%
2522	4.52	2574.0	6.12	1892.50	681.50	26.48%	6.00	1935.40	638.60	24.81%
2523	6.14	1115.0	6.12	1118.60	3.60	0.32%	6.00	1143.30	28.30	2.54%
2524	5.52	1525.0	6.12	1371.10	153.90	10.09%	6.00	1402.10	122.90	8.06%

Calibrated PROFAC: accurate to within 25%, 48% of the time.
Default PROFAC: accurate to within 25%, 52% of the time.

The mean PROFAC of calibration records, 6.12, is close to the default value of 6.00 for military ground applications (Table 3). As expected, the close mean and default PROFACs estimate effort with nearly the same accuracy; 48 versus 52% within 25%. The standard deviation of calibration PROFACs, 1.83, is 30% of the mean value of 6.12. The large standard deviation can be attributed to the PROFAC range of 2.27 to 9.74. Mean

relative error is slightly less for mean PROFAC estimated effort than it is for default PROFAC estimated effort; however, mean absolute error of mean PROFAC estimated effort is slightly *higher*. This is a recurring theme in this study and can be attributed to the fact that some records in the stratified data sets vary greatly by effort size. For example, one record may report actual effort of 5000, and the rest of the records may have an average actual effort of about 500. If the "more accurate PROFAC", as determined by the lower mean relative error, predicts the large records with less precision than normal, it can lead to a distortion of mean absolute errors; even though mean relative error (percentage error) remains relatively stable. In this example, records 55, 2219, 2520, 2522, and 2524 are large records that the mean calibrated PROFAC measured less accurately than the default PROFAC.

TABLE 4

MILITARY MOBILE

Military Mobile

Records Used for Calibration: 303, 347, 348, 349, 2456, 2502, 2507, 2508

Mean PROFAC of Calibration Records:	5.72
Standard Deviation:	2.16
Mean Absolute Error:	
Of Effort using Mean PROFAC	138.80
Of Effort using Default PROFAC	151.78
Mean Relative Error:	
Of Effort using Mean PROFAC	32.08%
Of Effort using Default PROFAC	33.34%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
303	8.59	250.0	5.72	370.30	120.30	48.12%	5.50	384.50	134.50	53.80%
347	3.25	41.0	5.72	23.30	17.70	43.17%	5.50	24.20	16.80	40.98%
348	4.13	418.0	5.72	304.30	113.70	27.20%	5.50	316.20	101.80	24.35%
349	4.55	59.0	5.72	47.20	11.80	20.00%	5.50	49.10	9.90	16.78%
2456	7.02	233.0	5.72	289.70	56.70	24.33%	5.50	302.10	69.10	29.66%
2502	8.46	743.6	5.72	1089.60	346.00	46.53%	5.50	1131.90	388.30	52.22%
2507	3.49	759.9	5.72	472.10	287.80	37.87%	5.50	490.50	269.40	35.45%
2508	6.30	1666.1	5.72	1822.50	156.40	9.39%	5.50	1890.50	224.40	13.47%

Calibrated PROFAC: accurate to within 25%, 38% of the time.
Default PROFAC: accurate to within 25%, 38% of the time.

Once again, the mean PROFAC of calibration records, 5.72, is close to the default value of 5.50 for military mobile applications (Table 4). Both mean PROFAC and default PROFAC estimate effort at an accuracy level of only 38% within 25%. For mean PROFAC estimated effort this is a attributable to the large PROFAC standard deviation of 2.16, which is 38% of the mean. Although the accuracy levels appear unpromising at first glance, the mean relative errors for both mean and default PROFAC estimated effort are only above the threshold accuracy level of 25% by 7 to 8%.

TABLE 5

MISSILE (1)

Missile (Iteration 1)

Records Used for Calibration: 16, 27, 36

Mean PROFAC of Calibration Records:	5.41
Standard Deviation:	0.44
Mean Absolute Error:	
Of Effort using Mean PROFAC	43.65
Of Effort using Default PROFAC	55.73
Mean Relative Error:	
Of Effort using Mean PROFAC	14.54%
Of Effort using Default PROFAC	15.18%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
15	3.08	276.0	5.41	163.00	113.00	40.94%	5.00	175.00	101.00	36.59%
16	5.86	96.0	5.41	104.10	8.10	8.44%	5.00	111.90	15.90	16.56%
27	5.37	1460.0	5.41	1446.00	14.00	0.96%	5.00	1563.50	103.50	7.09%
36	4.99	506.0	5.41	466.50	39.50	7.81%	5.00	503.50	2.50	0.49%

Calibrated PROFAC: accurate to within 25%, 75% of the time.
Default PROFAC: accurate to within 25%, 75% of the time.

For this iteration of the missile environmental platform, the mean PROFAC of calibration records, 5.41, is somewhat above the default value of 5.00 (Table 5). Regardless of that fact, both mean and default PROFAC estimated effort have an accuracy of 75% within 25%. The small standard deviation of 0.44 for calibrated PROFACs is due to record 15, a possible outlier, being absent from the computations during this iteration.

TABLE 6

MISSILE (2)

Missile (Iteration 2)

Records Used for Calibration: 15, 27, 36

Mean PROFAC of Calibration Records:	4.48
Standard Deviation:	1.23
Mean Absolute Error:	
Of Effort using Mean PROFAC	110.20
Of Effort using Default PROFAC	55.73
Mean Relative Error:	
Of Effort using Mean PROFAC	22.19%
Of Effort using Default PROFAC	15.18%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
15	3.08	276.0	4.48	194.60	81.40	29.49%	5.00	175.00	101.00	36.59%
16	5.86	96.0	4.48	124.70	28.70	29.90%	5.00	111.90	15.90	16.56%
27	5.37	1460.0	4.48	1738.80	278.80	19.10%	5.00	1563.50	103.50	7.09%
36	4.99	506.0	4.48	557.90	51.90	10.26%	5.00	503.50	2.50	0.49%

Calibrated PROFAC: accurate to within 25%, 50% of the time.
Default PROFAC: accurate to within 25%, 75% of the time.

With the inclusion of record 15 in the computations during this iteration, mean calibrated PROFAC plummets to 4.48 and the standard deviation nearly triples (Table 6). Mean relative error significantly increases from 14.54 to 22.19%. One less record of the four falls under the 25% accuracy threshold, but it should be noted that the two records outside of the 25% range have deviations of only 29.49 and 29.90%.

TABLE 7

MISSILE (3)

Missile (Iteration 3)	
Records Used for Calibration: 15, 16, 36	
Mean PROFAC of Calibration Records:	4.64
Standard Deviation:	1.42
Mean Absolute Error:	
Of Effort using Mean PROFAC	89.63
Of Effort using Default PROFAC	55.73
Mean Relative Error:	
Of Effort using Mean PROFAC	19.68%
Of Effort using Default PROFAC	15.18%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
15	3.08	276.0	4.64	188.60	87.40	31.67%	5.00	175.00	101.00	36.59%
16	5.86	96.0	4.64	120.70	24.70	25.73%	5.00	111.90	15.90	16.56%
27	5.37	1460.0	4.64	1672.00	212.00	14.52%	5.00	1563.50	103.50	7.09%
36	4.99	506.0	4.64	540.40	34.40	6.80%	5.00	503.50	2.50	0.49%

Calibrated PROFAC: accurate to within 25%, 50% of the time.
Default PROFAC: accurate to within 25%, 75% of the time.

This iteration brings a slight improvement to mean relative error of mean PROFAC estimated effort with the inclusion of record 16; which increases the mean calibrated PROFAC from 4.48 to 4.64 (Table 7).

TABLE 8

MISSILE (4)

Missile (Iteration 4)	
Records Used for Calibration: 15, 16, 27	
Mean PROFAC of Calibration Records:	4.77
Standard Deviation:	1.48
Mean Absolute Error:	
Of Effort using Mean PROFAC	76.60
Of Effort using Default PROFAC	55.73
Mean Relative Error:	
Of Effort using Mean PROFAC	17.89%
Of Effort using Default PROFAC	15.18%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
15	3.08	276.0	4.77	182.70	93.30	33.80%	5.00	175.00	101.00	36.59%
16	5.86	96.0	4.77	117.00	21.00	21.88%	5.00	111.90	15.90	16.56%
27	5.37	1460.0	4.77	1631.10	171.10	11.72%	5.00	1563.50	103.50	7.09%
36	4.99	506.0	4.77	527.00	21.00	4.15%	5.00	503.50	2.50	0.49%

Calibrated PROFAC: accurate to within 25%, 75% of the time.
Default PROFAC: accurate to within 25%, 75% of the time.

As the mean PROFAC continues to rise, mean relative error decreases further still; but not to the level of the first iteration (Table 8).

TABLE 9

UNMANNED SPACE

Unmanned Space

Records Used for Calibration: 38, 74, 75, 76, 78, 79, 81, 82, 305

Mean PROFAC of Calibration Records:	4.32
Standard Deviation:	1.89
Mean Absolute Error:	
Of Effort using Mean PROFAC	76.19
Of Effort using Default PROFAC	73.31
Mean Relative Error:	
Of Effort using Mean PROFAC	33.64%
Of Effort using Default PROFAC	33.55%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
29	4.21	52.0	4.32	50.60	1.40	2.69%	4.50	48.70	3.30	6.35%
38	2.66	64.0	4.32	39.50	24.50	38.28%	4.50	37.80	26.20	40.94%
74	3.03	80.0	4.32	56.70	23.30	29.13%	4.50	54.60	25.40	31.75%
75	5.56	912.0	4.32	1212.30	300.30	32.93%	4.50	1160.60	248.60	27.26%
76	3.29	115.0	4.32	88.30	26.70	23.22%	4.50	85.10	29.90	26.00%
77	2.48	523.0	4.32	301.70	221.30	42.31%	4.50	290.20	232.80	44.51%
78	4.46	478.0	4.32	495.20	17.20	3.60%	4.50	473.70	4.30	0.90%
79	3.67	432.0	4.32	367.10	64.90	15.02%	4.50	353.00	79.00	18.29%
80	3.02	296.0	4.32	206.90	89.10	30.10%	4.50	199.00	97.00	32.77%
81	4.70	164.0	4.32	178.10	14.10	8.60%	4.50	171.70	7.70	4.70%
82	2.86	140.0	4.32	94.10	45.90	32.79%	4.50	90.60	49.40	35.29%
83	3.35	57.0	4.32	44.50	12.50	21.93%	4.50	42.80	14.20	24.91%
305	8.66	145.0	4.32	287.70	142.70	98.41%	4.50	277.10	132.10	91.10%
306	8.41	90.0	4.32	172.80	82.80	92.00%	4.50	166.40	76.40	84.89%

Calibrated PROFAC: accurate to within 25%, 43% of the time.
Default PROFAC: accurate to within 25%, 36% of the time.

The mean PROFAC of calibration records, 4.32, is close to the default value of 4.50 for unmanned space applications (Table 9). As expected, the close mean and default PROFACs estimate effort with nearly the same accuracy; 43 versus 36% within 25%. The standard deviation of calibration PROFACs, 1.89, is 44% of the mean value of 4.32. The large standard deviation can be attributed to the PROFAC range of 2.48 to 8.66. Once again it should be noted that mean relative errors for both mean and default estimated efforts are only about 8% above the threshold accuracy level of 25%.

TABLE 10

ADA

ADA

Records Used for Calibration: 29, 347, 348, 349, 2501, 2502, 2508, 2512

Mean PROFAC of Calibration Records:	5.51
Standard Deviation:	1.72
Mean Absolute Error:	
Of Effort using Mean PROFAC	116.63
Of Effort using Default PROFAC	90.45
Mean Relative Error:	
Of Effort using Mean PROFAC	28.16%
Of Effort using Default PROFAC	23.57%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
29	4.25	52.0	5.51	40.60	11.40	21.92%	5.00	44.60	7.40	14.23%
347	3.30	41.0	5.51	24.50	16.50	40.24%	6.00	22.50	18.50	45.12%
348	4.13	418.0	5.51	323.60	94.40	22.58%	6.00	299.50	118.50	28.35%
349	4.60	59.0	5.51	49.30	9.70	16.44%	6.00	45.50	13.50	22.88%
2497	7.96	80.0	5.51	113.50	33.50	41.88%	6.50	97.00	17.00	21.25%
2501	6.78	418.0	5.51	520.50	102.50	24.52%	6.50	437.40	19.40	4.64%
2502	8.46	743.6	5.51	1118.60	375.00	50.43%	6.00	1032.20	288.60	38.81%
2507	3.50	759.9	5.51	493.50	266.40	35.06%	6.00	454.70	305.20	40.16%
2508	6.30	1666.1	5.51	1882.80	216.70	13.01%	6.00	1741.70	75.60	4.54%
2512	6.25	259.0	5.51	299.20	40.20	15.52%	5.50	299.80	40.80	15.75%

Calibrated PROFAC: accurate to within 25%, 60% of the time.
Default PROFAC: accurate to within 25%, 60% of the time.

The mean PROFAC of calibration records, 5.51, is somewhat lower than the average default PROFAC of 5.95 for these Ada records (Table 10). The mean and default PROFACs estimate effort with the same accuracy; 60% within 25%; however, mean relative error of default PROFAC estimated effort is nearly 5% lower than the error of mean PROFAC estimated effort. The standard deviation of calibration PROFACs, 1.72, is 31% of the mean value of 5.51. The large standard deviation can be attributed to the PROFAC range of 3.30 to 8.46.

TABLE 11

ASSEMBLY

ASSEMBLY

Records Used for Calibration: 15, 16, 36, 53, 55, 56, 61, 302

Mean PROFAC of Calibration Records:	6.10
Standard Deviation:	2.51
Mean Absolute Error:	
Of Effort using Mean PROFAC	469.69
Of Effort using Default PROFAC	439.63
Mean Relative Error:	
Of Effort using Mean PROFAC	29.67%
Of Effort using Default PROFAC	32.29%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
15	3.08	276.0	6.10	145.10	130.90	47.43%	4.50	193.80	82.20	29.78%
16	5.86	96.0	6.10	92.60	3.40	3.54%	4.50	124.20	28.20	29.38%
36	4.99	506.0	6.10	414.90	91.10	18.00%	4.50	555.60	49.60	9.80%
38	2.66	64.0	6.10	27.90	36.10	56.41%	4.00	42.60	21.40	33.44%
53	7.51	51.0	6.10	62.70	11.70	22.94%	5.50	69.60	18.60	36.47%
55	2.27	6713.0	6.10	2398.80	4314.20	64.27%	5.50	2685.80	4027.20	59.99%
56	7.34	23.0	6.10	28.00	5.00	21.74%	5.50	31.40	8.40	36.52%
61	9.20	113.0	6.10	167.10	54.10	47.88%	5.50	182.30	69.30	61.33%
302	8.55	400.0	6.10	437.60	37.60	9.40%	4.50	469.10	69.10	17.28%
303	7.70	250.0	6.10	262.80	12.80	5.12%	5.00	272.30	22.30	8.92%

Calibrated PROFAC: accurate to within 25%, 60% of the time.
Default PROFAC: accurate to within 25%, 30% of the time.

The mean PROFAC of calibration records, 6.10, is much higher than the average default PROFAC of 4.90 for these Assembly records (Table 11). The mean PROFAC estimates effort with greater accuracy than the default values; 60 versus 30% within 25%. Mean absolute error of mean PROFAC estimated effort is distorted due to the influence of record 55; which has a calibrated PROFAC of only 2.27 and an actual effort of 6713.

TABLE 12

C (1)

C (Iteration1)

Records Used for Calibration: 306, 307, 2510

Mean PROFAC of Calibration Records:	8.57
Standard Deviation:	0.26
Mean Absolute Error:	
Of Effort using Mean PROFAC	1.93
Of Effort using Default PROFAC	48.25
Mean Relative Error:	
Of Effort using Mean PROFAC	1.98%
Of Effort using Default PROFAC	43.85%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
305	8.66	145.0	8.57	146.10	1.10	0.76%	5.00	205.80	60.80	41.93%
306	8.41	90.0	8.57	88.40	1.60	1.78%	5.00	149.90	59.90	66.56%
307	8.87	52.0	8.57	53.90	1.90	3.65%	6.50	71.70	19.70	37.88%
2510	8.42	181.2	8.57	178.10	3.10	1.71%	6.50	233.80	52.60	29.03%

Calibrated PROFAC: accurate to within 25%, 100% of the time.
Default PROFAC: accurate to within 25%, 0% of the time.

The mean PROFAC of calibration records, 8.57, again is much higher than the average default PROFAC of 5.75 for these C records (Table 12). The mean PROFAC estimates effort with an outstanding accuracy level of 100% within 25%, with a mean relative error of only 1.98%; while none of the default estimated efforts are within 25% of actuals. The mean PROFAC estimates are extremely close to actuals because the standard deviation of the mean is only 0.26.

TABLE 13

C (2)

C (Iteration 2)

Records Used for Calibration: 305, 307, 2510

Mean PROFAC of Calibration Records:	8.65
Standard Deviation:	0.23
Mean Absolute Error:	
Of Effort using Mean PROFAC	2.13
Of Effort using Default PROFAC	48.25
Mean Relative Error:	
Of Effort using Mean PROFAC	1.99%
Of Effort using Default PROFAC	43.85%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
305	8.66	145.0	8.65	145.10	0.10	0.07%	5.00	205.80	60.80	41.93%
306	8.41	90.0	8.65	87.60	2.40	2.67%	5.00	149.90	59.90	66.56%
307	8.87	52.0	8.65	53.40	1.40	2.69%	6.50	71.70	19.70	37.88%
2510	8.42	181.2	8.65	176.60	4.60	2.54%	6.50	233.80	52.60	29.03%

Calibrated PROFAC: accurate to within 25%, 100% of the time.
Default PROFAC: accurate to within 25%, 0% of the time.

This iteration is nearly identical to the first one. Record 305 is substituted for 306 in the computations (Table 13).

TABLE 14

C (3)

C (Iteration 3)

Records Used for Calibration: 305, 306, 2510

Mean PROFAC of Calibration Records:	8.50
Standard Deviation:	0.14
Mean Absolute Error:	
Of Effort using Mean PROFAC	1.70
Of Effort using Default PROFAC	48.25
Mean Relative Error:	
Of Effort using Mean PROFAC	1.92%
Of Effort using Default PROFAC	43.85%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
305	8.66	145.0	8.50	146.90	1.90	1.31%	5.00	205.80	60.80	41.93%
306	8.41	90.0	8.50	89.10	0.90	1.00%	5.00	149.90	59.90	66.56%
307	8.87	52.0	8.50	54.30	2.30	4.42%	6.50	71.70	19.70	37.88%
2510	8.42	181.2	8.50	179.50	1.70	0.94%	6.50	233.80	52.60	29.03%

Calibrated PROFAC: accurate to within 25%, 100% of the time.

Default PROFAC: accurate to within 25%, 0% of the time.

In this iteration, standard deviation drops to a low 0.14 as the mean PROFAC decreases to 8.50 (Table 14). Mean relative error falls to 1.92%.

TABLE 15

C (4)

C (Iteration 4)

Records Used for Calibration: 305, 306, 307

Mean PROFAC of Calibration Records:	8.65
Standard Deviation:	0.23
Mean Absolute Error:	
Of Effort using Mean PROFAC	2.13
Of Effort using Default PROFAC	48.25
Mean Relative Error:	
Of Effort using Mean PROFAC	1.99%
Of Effort using Default PROFAC	43.85%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
305	8.66	145.0	8.65	145.10	0.10	0.07%	5.00	205.80	60.80	41.93%
306	8.41	90.0	8.65	87.60	2.40	2.67%	5.00	149.90	59.90	66.56%
307	8.87	52.0	8.65	53.40	1.40	2.69%	6.50	71.70	19.70	37.88%
2510	8.42	181.2	8.65	176.60	4.60	2.54%	6.50	233.80	52.60	29.03%

Calibrated PROFAC: accurate to within 25%, 100% of the time.
Default PROFAC: accurate to within 25%, 0% of the time.

The results of this iteration are identical to the second iteration (Table 15). Record 306 is substituted for record 2510. The calibrated PROFACs of the records differ by only 0.01.

TABLE 16

COBOL (1)

COBOL (Iteration 1)

Records Used for Calibration: 2520, 2521, 2522, 2523, 2524

Mean PROFAC of Calibration Records:	4.39
Standard Deviation:	1.55
Mean Absolute Error:	
Of Effort using Mean PROFAC	780.72
Of Effort using Default PROFAC	744.67
Mean Relative Error:	
Of Effort using Mean PROFAC	39.66%
Of Effort using Default PROFAC	32.19%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
2519	8.22	1833.0	4.39	3672.40	1839.40	100.35%	6.50	2388.20	555.20	30.29%
2520	2.44	3960.0	4.39	2201.60	1758.40	44.40%	6.50	1486.00	2474.00	62.47%
2521	3.27	735.0	4.39	545.00	190.00	25.85%	6.50	366.50	368.50	50.14%
2522	4.52	2574.0	4.39	2644.80	70.80	2.75%	6.50	1784.60	789.40	30.67%
2523	6.20	1115.0	4.39	1549.40	434.40	38.96%	6.50	1066.20	48.80	4.38%
2524	5.52	1525.0	4.39	1916.30	391.30	25.66%	6.50	1292.90	232.10	15.22%

Calibrated PROFAC: accurate to within 25%, 17% of the time.
Default PROFAC: accurate to within 25%, 33% of the time.

The mean PROFAC of calibration records, 4.39, is much lower than the average default PROFAC of 6.50 for these COBOL records (Table 16). The mean and default PROFACs both estimate effort with poor accuracy of 17 versus 33% within 25%; respectively. Calibrated PROFACs range from 2.44 to 8.22.

TABLE 17

COBOL (2)

COBOL (Iteration 2)

Records Used for Calibration: 2519, 2521, 2522, 2523, 2524

Mean PROFAC of Calibration Records:	5.55
Standard Deviation:	1.86
Mean Absolute Error:	
Of Effort using Mean PROFAC	692.12
Of Effort using Default PROFAC	744.67
Mean Relative Error:	
Of Effort using Mean PROFAC	30.59%
Of Effort using Default PROFAC	32.19%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
2519	8.22	1833.0	5.55	2847.50	1014.50	55.35%	6.50	2388.20	555.20	30.29%
2520	2.44	3960.0	5.55	1743.80	2216.20	55.96%	6.50	1486.00	2474.00	62.47%
2521	3.27	735.0	5.55	430.70	304.30	41.40%	6.50	366.50	368.50	50.14%
2522	4.52	2574.0	5.55	2094.50	479.50	18.63%	6.50	1784.60	789.40	30.67%
2523	6.20	1115.0	5.55	1245.60	130.60	11.71%	6.50	1066.20	48.80	4.38%
2524	5.52	1525.0	5.55	1517.40	7.60	0.50%	6.50	1292.90	232.10	15.22%

Calibrated PROFAC: accurate to within 25%, 50% of the time.
Default PROFAC: accurate to within 25%, 33% of the time.

Mean relative error decreases dramatically from 39.66 to 30.59 as mean calibrated PROFAC increases from 4.39 to 5.55 (Table 17). Mean PROFAC estimated effort increases in accuracy to 50% within 25%.

TABLE 18

COBOL (3)

COBOL (Iteration 3)

Records Used for Calibration: 2519, 2520, 2522, 2523, 2524

Mean PROFAC of Calibration Records:	5.38
Standard Deviation:	2.13
Mean Absolute Error:	
Of Effort using Mean PROFAC	697.58
Of Effort using Default PROFAC	744.67
Mean Relative Error:	
Of Effort using Mean PROFAC	31.41%
Of Effort using Default PROFAC	32.19%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
2519	8.22	1833.0	5.38	2943.30	1110.30	60.57%	6.50	2388.20	555.20	30.29%
2520	2.44	3960.0	5.38	1797.00	2163.00	54.62%	6.50	1486.00	2474.00	62.47%
2521	3.27	735.0	5.38	444.00	291.00	39.59%	6.50	366.50	368.50	50.14%
2522	4.52	2574.0	5.38	2158.40	415.60	16.15%	6.50	1784.60	789.40	30.67%
2523	6.20	1115.0	5.38	1281.80	166.80	14.96%	6.50	1066.20	48.80	4.38%
2524	5.52	1525.0	5.38	1563.80	38.80	2.54%	6.50	1292.90	232.10	15.22%

Calibrated PROFAC: accurate to within 25%, 50% of the time.
Default PROFAC: accurate to within 25%, 33% of the time.

This iteration introduces slightly more relative error than the previous iteration as mean PROFAC decreases to 5.38 (Table 18).

TABLE 19

COBOL (4)

COBOL (Iteration 4)

Records Used for Calibration: 2519, 2520, 2521, 2523, 2524

Mean PROFAC of Calibration Records:	5.13
Standard Deviation:	2.32
Mean Absolute Error:	
Of Effort using Mean PROFAC	710.05
Of Effort using Default PROFAC	744.67
Mean Relative Error:	
Of Effort using Mean PROFAC	32.95%
Of Effort using Default PROFAC	32.19%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
2519	8.22	1833.0	5.13	3095.60	1262.60	68.88%	6.50	2388.20	555.20	30.29%
2520	2.44	3960.0	5.13	1882.00	2078.00	52.47%	6.50	1486.00	2474.00	62.47%
2521	3.27	735.0	5.13	465.20	269.80	36.71%	6.50	366.50	368.50	50.14%
2522	4.52	2574.0	5.13	2260.70	313.30	12.17%	6.50	1784.60	789.40	30.67%
2523	6.20	1115.0	5.13	1338.70	223.70	20.06%	6.50	1066.20	48.80	4.38%
2524	5.52	1525.0	5.13	1637.90	112.90	7.40%	6.50	1292.90	232.10	15.22%

Calibrated PROFAC: accurate to within 25%, 50% of the time.
Default PROFAC: accurate to within 25%, 33% of the time.

This iteration is similar to the previous one. Mean PROFAC decreases to 5.13 (Table 19). Records 2520, 2521, 2522, and 2524 distort values for mean absolute error.

TABLE 20

COBOL (5)

COBOL (Iteration 5)

Records Used for Calibration: 2519, 2520, 2521, 2522, 2524

Mean PROFAC of Calibration Records:	4.79
Standard Deviation:	2.25
Mean Absolute Error:	
Of Effort using Mean PROFAC	731.08
Of Effort using Default PROFAC	744.67
Mean Relative Error:	
Of Effort using Mean PROFAC	35.54%
Of Effort using Default PROFAC	32.19%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
2519	8.22	1833.0	4.79	3351.80	1518.80	82.86%	6.50	2388.20	555.20	30.29%
2520	2.44	3960.0	4.79	2024.40	1935.60	48.88%	6.50	1486.00	2474.00	62.47%
2521	3.27	735.0	4.79	500.70	234.30	31.88%	6.50	366.50	368.50	50.14%
2522	4.52	2574.0	4.79	2431.80	142.20	5.52%	6.50	1784.60	789.40	30.67%
2523	6.20	1115.0	4.79	1433.80	318.80	28.59%	6.50	1066.20	48.80	4.38%
2524	5.52	1525.0	4.79	1761.80	236.80	15.53%	6.50	1292.90	232.10	15.22%

Calibrated PROFAC: accurate to within 25%, 33% of the time.
Default PROFAC: accurate to within 25%, 33% of the time.

As the mean PROFAC continues to drop, error increases still further (Table 20).

TABLE 21

COBOL (6)

COBOL (Iteration 6)

Records Used for Calibration: 2519, 2520, 2521, 2522, 2523

Mean PROFAC of Calibration Records:	4.93
Standard Deviation:	2.32
Mean Absolute Error:	
Of Effort using Mean PROFAC	721.32
Of Effort using Default PROFAC	744.67
Mean Relative Error:	
Of Effort using Mean PROFAC	34.40%
Of Effort using Default PROFAC	32.19%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
2519	8.22	1833.0	4.93	3250.80	1417.80	77.35%	6.50	2388.20	555.20	30.29%
2520	2.44	3960.0	4.93	1968.90	1991.10	50.28%	6.50	1486.00	2474.00	62.47%
2521	3.27	735.0	4.93	486.90	248.10	33.76%	6.50	366.50	368.50	50.14%
2522	4.52	2574.0	4.93	2365.10	208.90	8.12%	6.50	1784.60	789.40	30.67%
2523	6.20	1115.0	4.93	1388.50	273.50	24.53%	6.50	1066.20	48.80	4.38%
2524	5.52	1525.0	4.93	1713.50	188.50	12.36%	6.50	1292.90	232.10	15.22%

Calibrated PROFAC: accurate to within 25%, 50% of the time.
Default PROFAC: accurate to within 25%, 33% of the time.

This iteration is similar to the previous one. The mean PROFAC is slightly higher (Table 21).

TABLE 22

FORTRAN

FORTRAN

Records Used for Calibration: 2, 7, 9, 48, 63, 68, 80, 2456

Mean PROFAC of Calibration Records:	5.62
Standard Deviation:	1.86
Mean Absolute Error:	
Of Effort using Mean PROFAC	72.94
Of Effort using Default PROFAC	57.08
Mean Relative Error:	
Of Effort using Mean PROFAC	26.50%
Of Effort using Default PROFAC	19.74%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
2	3.25	245.0	5.62	137.50	107.50	43.88%	6.50	117.60	127.40	52.00%
7	6.80	127.0	5.62	150.10	23.10	18.19%	6.50	131.70	4.70	3.70%
9	6.75	545.0	5.62	657.80	112.80	20.70%	6.50	565.80	20.80	3.82%
48	3.50	344.0	5.62	197.80	146.20	42.50%	6.50	166.60	177.40	51.57%
50	6.35	656.0	5.62	753.20	97.20	14.82%	6.50	639.60	16.40	2.50%
63	7.19	23.0	5.62	31.20	8.20	35.65%	6.50	26.10	3.10	13.48%
68	7.00	170.0	5.62	212.90	42.90	25.24%	6.50	183.70	13.70	8.06%
80	3.40	296.0	5.62	219.00	77.00	26.01%	5.00	236.80	59.20	20.00%
301	4.90	430.0	5.62	373.60	56.40	13.12%	6.50	321.90	108.10	25.14%
2456	7.05	233.0	5.62	291.10	58.10	24.94%	6.00	273.00	40.00	17.17%

Calibrated PROFAC: accurate to within 25%, 50% of the time.
Default PROFAC: accurate to within 25%, 70% of the time.

The mean PROFAC of calibration records, 5.62, is much lower than the average default PROFAC of 6.30 for these FORTRAN records (Table 22). The default PROFACs estimate effort with greater accuracy than the mean PROFAC; 70 versus 50% within 25%. It should be noted, however, that the mean relative error of mean PROFAC estimated effort is only 1.50% above the accuracy threshold of 25%. Two of the records fall just outside the borderline and lead to only 50% of the records being estimated within 25%.

TABLE 23

JOVIAL

JOVIAL

Records Used for Calibration: 14, 27, 74, 77, 78, 80, 81, 302

Mean PROFAC of Calibration Records:	4.33
Standard Deviation:	1.96
Mean Absolute Error:	
Of Effort using Mean PROFAC	113.91
Of Effort using Default PROFAC	115.00
Mean Relative Error:	
Of Effort using Mean PROFAC	26.15%
Of Effort using Default PROFAC	25.05%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
14	3.44	490.0	4.33	393.30	96.70	19.73%	4.50	379.60	110.40	22.53%
27	5.37	1460.0	4.33	1792.70	332.70	22.79%	4.50	1731.90	271.90	18.62%
74	3.03	80.0	4.33	56.50	23.50	29.38%	4.00	61.30	18.70	23.38%
75	5.56	912.0	4.33	1209.30	297.30	32.60%	4.00	1315.50	403.50	44.24%
76	3.29	115.0	4.33	88.10	26.90	23.39%	4.00	94.70	20.30	17.65%
77	2.48	523.0	4.33	301.00	222.00	42.45%	4.00	324.70	198.30	37.92%
78	4.46	478.0	4.33	494.00	16.00	3.35%	4.00	538.50	60.50	12.66%
79	3.67	432.0	4.33	366.30	65.70	15.21%	4.00	395.20	36.80	8.52%
80	2.70	296.0	4.33	230.20	65.80	22.23%	4.00	240.60	55.40	18.72%
81	4.70	164.0	4.33	177.70	13.70	8.35%	4.00	190.90	26.90	16.40%
82	2.86	140.0	4.33	93.90	46.10	32.93%	4.00	101.00	39.00	27.86%
83	3.35	57.0	4.33	44.40	12.60	22.11%	4.00	47.80	9.20	16.14%
302	8.45	400.0	4.33	661.80	261.80	65.45%	4.50	644.10	244.10	61.03%

Calibrated PROFAC: accurate to within 25%, 62% of the time.
Default PROFAC: accurate to within 25%, 69% of the time.

The mean PROFAC of calibration records, 4.33, is slightly higher than the average default PROFAC of 4.12 for these JOVIAL records (Table 23). As expected, the mean and default PROFACs estimate effort with similar accuracy; 62 versus 69% within 25%; respectively. Those are surprising accuracy levels considering that the standard deviation of 1.96 is 45% of the mean. Record 75 causes the largest distortion to mean absolute error values.

TABLE 24

PASCAL (1)

PASCAL (Iteration 1)

Records Used for Calibration: 54, 70, 72, 73, 303

Mean PROFAC of Calibration Records:	8.18
Standard Deviation:	1.71
Mean Absolute Error:	
Of Effort using Mean PROFAC	24.43
Of Effort using Default PROFAC	43.27
Mean Relative Error:	
Of Effort using Mean PROFAC	20.89%
Of Effort using Default PROFAC	35.14%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
30	4.10	160.0	8.18	79.60	80.40	50.25%	6.50	100.90	59.10	36.94%
54	9.80	134.0	8.18	165.90	31.90	23.81%	6.50	218.30	84.30	62.91%
70	5.30	40.0	8.18	25.00	15.00	37.50%	6.50	31.90	8.10	20.25%
72	8.70	42.0	8.18	45.00	3.00	7.14%	6.50	55.90	13.90	33.10%
73	8.20	56.0	8.18	56.10	0.10	0.18%	6.50	70.40	14.40	25.71%
303	8.90	250.0	8.18	266.20	16.20	6.48%	6.00	329.80	79.80	31.92%

Calibrated PROFAC: accurate to within 25%, 67% of the time.
Default PROFAC: accurate to within 25%, 17% of the time.

The mean PROFAC of calibration records, 8.18, is much higher than the average default PROFAC of 6.42 for these PASCAL records (Table 24). The mean PROFAC estimates effort with much greater accuracy than the default values; 67 versus 17% within 25%; respectively. This is due to the fact that the standard deviation of 1.71 is only 21% of the mean calibrated PROFAC and the mean PROFAC is much larger than the average default value.

TABLE 25

PASCAL (2)

PASCAL (Iteration 2)

Records Used for Calibration: 30, 70, 72, 73, 303

Mean PROFAC of Calibration Records:	7.04
Standard Deviation:	2.19
Mean Absolute Error:	
Of Effort using Mean PROFAC	34.50
Of Effort using Default PROFAC	43.27
Mean Relative Error:	
Of Effort using Mean PROFAC	29.11%
Of Effort using Default PROFAC	35.14%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
30	4.10	160.0	7.04	92.80	67.20	42.00%	6.50	100.90	59.10	36.94%
54	9.80	134.0	7.04	198.20	64.20	47.91%	6.50	218.30	84.30	62.91%
70	5.30	40.0	7.04	29.20	10.80	27.00%	6.50	31.90	8.10	20.25%
72	8.70	42.0	7.04	51.70	9.70	23.10%	6.50	55.90	13.90	33.10%
73	8.20	56.0	7.04	65.10	9.10	16.25%	6.50	70.40	14.40	25.71%
303	8.90	250.0	7.04	296.00	46.00	18.40%	6.00	329.80	79.80	31.92%

Calibrated PROFAC: accurate to within 25%, 50% of the time.
Default PROFAC: accurate to within 25%, 17% of the time.

Mean relative error increases dramatically as record 30 is included in the computations. Record 30 decreases the mean PROFAC from 8.18 to 7.04 (Table 25).

TABLE 26

PASCAL (3)

PASCAL (Iteration 3)

Records Used for Calibration: 30, 54, 72, 73, 303

Mean PROFAC of Calibration Records:	7.94
Standard Deviation:	2.22
Mean Absolute Error:	
Of Effort using Mean PROFAC	26.50
Of Effort using Default PROFAC	43.27
Mean Relative Error:	
Of Effort using Mean PROFAC	22.57%
Of Effort using Default PROFAC	35.14%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
30	4.10	160.0	7.94	82.00	78.00	48.75%	6.50	100.90	59.10	36.94%
54	9.80	134.0	7.94	172.80	38.80	28.96%	6.50	218.30	84.30	62.91%
70	5.30	40.0	7.94	25.80	14.20	35.50%	6.50	31.90	8.10	20.25%
72	8.70	42.0	7.94	46.30	4.30	10.24%	6.50	55.90	13.90	33.10%
73	8.20	56.0	7.94	57.80	1.80	3.21%	6.50	70.40	14.40	25.71%
303	8.90	250.0	7.94	271.90	21.90	8.76%	6.00	329.80	79.80	31.92%

Calibrated PROFAC: accurate to within 25%, 50% of the time.
Default PROFAC: accurate to within 25%, 17% of the time.

This iteration is similar to iteration one. It includes only one of the possible outliers (record 30) in the computations (Table 26).

TABLE 27

PASCAL (4)

PASCAL (Iteration 4)

Records Used for Calibration: 30, 54, 70, 73, 303

Mean PROFAC of Calibration Records:	7.26
Standard Deviation:	2.44
Mean Absolute Error:	
Of Effort using Mean PROFAC	32.37
Of Effort using Default PROFAC	43.27
Mean Relative Error:	
Of Effort using Mean PROFAC	27.31%
Of Effort using Default PROFAC	35.14%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
30	4.10	160.0	7.26	89.80	70.20	43.88%	6.50	100.90	59.10	36.94%
54	9.80	134.0	7.26	191.40	57.40	42.84%	6.50	218.30	84.30	62.91%
70	5.30	40.0	7.26	28.30	11.70	29.25%	6.50	31.90	8.10	20.25%
72	8.70	42.0	7.26	50.20	8.20	19.52%	6.50	55.90	13.90	33.10%
73	8.20	56.0	7.26	63.00	7.00	12.50%	6.50	70.40	14.40	25.71%
303	8.90	250.0	7.26	289.70	39.70	15.88%	6.00	329.80	79.80	31.92%

Calibrated PROFAC: accurate to within 25%, 50% of the time.
Default PROFAC: accurate to within 25%, 17% of the time.

These final three iterations are all similar to iteration two in that both possible outliers (records 30 and 70) are included in the computations. Here, mean PROFAC decreases to 7.26; increasing mean relative error to 27.31% (Table 27).

TABLE 28

PASCAL (5)

PASCAL (Iteration 5)

Records Used for Calibration: 30, 54, 70, 72, 303

Mean PROFAC of Calibration Records:	7.36
Standard Deviation:	2.50
Mean Absolute Error:	
Of Effort using Mean PROFAC	31.45
Of Effort using Default PROFAC	43.27
Mean Relative Error:	
Of Effort using Mean PROFAC	26.55%
Of Effort using Default PROFAC	35.14%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
30	4.10	160.0	7.36	88.50	71.50	44.69%	6.50	100.90	59.10	36.94%
54	9.80	134.0	7.36	188.40	54.40	40.60%	6.50	218.30	84.30	62.91%
70	5.30	40.0	7.36	27.90	12.10	30.25%	6.50	31.90	8.10	20.25%
72	8.70	42.0	7.36	49.60	7.60	18.10%	6.50	55.90	13.90	33.10%
73	8.20	56.0	7.36	62.10	6.10	10.89%	6.50	70.40	14.40	25.71%
303	8.90	250.0	7.36	287.00	37.00	14.80%	6.00	329.80	79.80	31.92%

Calibrated PROFAC: accurate to within 25%, 50% of the time.
Default PROFAC: accurate to within 25%, 17% of the time.

Mean PROFAC increases slightly with the substitution of record 72 for 73 (Table 28).

TABLE 29

PASCAL (6)

PASCAL (Iteration 6)

Records Used for Calibration: 30, 54, 70, 72, 73

Mean PROFAC of Calibration Records:	7.22
Standard Deviation:	2.41
Mean Absolute Error:	
Of Effort using Mean PROFAC	32.77
Of Effort using Default PROFAC	43.27
Mean Relative Error:	
Of Effort using Mean PROFAC	27.65%
Of Effort using Default PROFAC	35.14%

Record #	Cal PROFAC	Actual Effort	Mean PROFAC	Effort	Dev. from Actual	% Dev.	Def PROFAC	Effort	Dev. from Actual	% Dev.
30	4.10	160.0	7.22	90.30	69.70	43.56%	6.50	100.90	59.10	36.94%
54	9.80	134.0	7.22	192.60	58.60	43.73%	6.50	218.30	84.30	62.91%
70	5.30	40.0	7.22	28.40	11.60	29.00%	6.50	31.90	8.10	20.25%
72	8.70	42.0	7.22	50.50	8.50	20.24%	6.50	55.90	13.90	33.10%
73	8.20	56.0	7.22	63.30	7.30	13.04%	6.50	70.40	14.40	25.71%
303	8.90	250.0	7.22	290.90	40.90	16.36%	6.00	329.80	79.80	31.92%

Calibrated PROFAC: accurate to within 25%, 50% of the time.
Default PROFAC: accurate to within 25%, 17% of the time.

Mean PROFAC drops again as record 73 is substituted for 303. Accuracy of estimated effort remains at 50% within 25% (Table 29).

Summary

Twenty-seven PRICE S calibration iterations concentrating on operating environment and language were performed using data records from the SMC Database. The results were surprising in that, in most instances, the mean calibrated PROFAC values estimated actual effort well, but not overwhelmingly better than the default PROFAC values. The main contributing factor to this phenomena was variability within the stratified data sets.

V. Conclusions and Recommendations for Follow-on Research

Conclusions

The calibration effort was encouraging in that the results from seven of the eleven stratified data sets suggested either a new refinement in PROFAC values based upon the mean calibrated PROFAC or the recommendation to use PROFAC values from analogous calibrated records for estimating future efforts. In most cases, however, the calibrated PROFAC estimated effort only slightly better than the default values. A summary of the conclusions can be found in Table 30, and is followed by a detailed description of the conclusions for each environmental platform or language.

TABLE 30

PROFAC RECOMMENDATION SUMMARY

Data Set	Mean Cal PROFAC	MRE	Mean Def PROFAC	MRE	PROFAC Recommendation
Military Ground	6.12	29.41%	6.00	30.33%	Analogous Records
Military Mobile	5.72	32.08%	5.50	33.34%	Analogous Records
Missile	5.41	14.54%	5.00	15.18%	Default
Unmanned Space	4.32	33.64%	4.50	33.55%	Analogous Records
Ada	5.51	28.16%	5.95	23.57%	Default or Analogous Records
Assembly	6.10	29.67%	4.90	32.29%	Calibrated or Analogous Records
C	8.50	1.92%	5.75	43.85%	Calibrated
COBOL	5.55	30.59%	6.50	32.19%	Analogous Records
FORTRAN	5.62	26.50%	6.30	19.74%	Default or Analogous Records
JOVIAL	4.33	26.15%	4.12	25.05%	Default or Analogous Records
PASCAL	8.18	20.89%	6.42	35.14%	Calibrated or Analogous Records

Military Ground

The mean calibrated PROFAC and the default PROFAC were close with only a difference of 0.12. Neither PROFAC, however, measured effort consistently in the study as evidenced by the percentage within 25% of 48% for calibrated and 52% for default. This poor accuracy level was caused by the 2.27 to 9.74 range of calibrated PROFAC values. With such variation, the recommendation is to find analogous records within the database when considering a value for PROFAC in a new development effort. A single PROFAC applied across all military ground programs is ill-advised.

Military Mobile

Once again, the mean calibrated PROFAC and default PROFAC values were very close and both only measured effort within 25%, 38% of the time. The variability in the data set leads to the recommendation to use analogous records for PROFAC for new programs.

Missile

In this data set, both the calibrated and default PROFACs measured effort within 25%, 75% of the time during the most accurate iteration. Both had very low mean relative error values; 14.54 and 15.18% for calibrated and default estimates, respectively. However, the default PROFAC value of 5.00 yielded lower mean relative errors for three of the four iterations. The recommendation is to continue to use of the default value of 5.00.

Unmanned Space

This environmental platform yielded comparable results to the Military Ground and Military Mobile platforms. Huge variability within the data set led to poor estimating accuracy for both calibrated and default PROFAC values. This may be attributable in part to the suspicion that some of the Unmanned Space records in the database should actually be classified as Military Ground. The use of analogous records for future efforts is recommended.

Ada

Default values for Ada ranged from 5.00 to 6.50 with a mean of 5.95. Both the default values and the single mean calibrated PROFAC estimated effort with similar accuracy; 60% within 25%. The default estimations, however, had a lower value for mean relative error; 23.57 versus 28.16%. The recommendation is to use analogous records to achieve the greatest accuracy possible for new estimates; however, the default values should be used in the absence of well-defined new project parameters.

Assembly

The mean calibrated PROFAC, 6.10, was significantly greater than the mean default value of 4.90. Default PROFAC values ranged from 4.50 to 5.50 in this data set. The mean calibrated PROFAC estimated effort much more accurately than the default values; 60 versus 30% within 25%. The large mean relative error of 29.67% and the standard deviation of 2.51 for mean calibrated PROFAC, however, precludes a recommendation of solely relying on the calibrated value for PROFAC. Analogous records should be considered first when estimating new development efforts.

C

All four C iterations yielded an accuracy level of 100% within 25% using mean calibrated PROFAC values. Default values ranged from 5.00 to 6.50 and estimated within 25%, 0% of the time. The most accurate iteration had a mean PROFAC value of 8.50 and a standard deviation of only 0.14. The recommendation is to use the calibrated value of 8.50 for future estimates.

COBOL

The mean calibrated PROFAC, 5.55, was drastically different from the default value of 6.50. The mean calibrated PROFAC estimated effort within 25%, 50% of the time versus 33% for default estimates during the most accurate iteration. Both calibrated and default estimates, however, had mean relative errors above 30%. The recommendation for future estimates is to research analogous records.

FORTRAN

Once again, the mean calibrated and default PROFACs were drastically different; 5.62 versus 6.30. Default values ranged from 5.00 to 6.50. The default values estimated effort within 25%, 70% of the time versus 50% for calibrated estimates. Although the default estimates had a low mean relative error of 19.74%, there was noticeable variation in the estimates as two of the records were estimated with errors over 50%. The recommendation is to use analogous records to achieve the greatest accuracy possible for new estimates; however, the default values should be used in the absence of well-defined new project parameters.

JOVIAL

The mean calibrated PROFAC, 4.33, was very close to the mean default value of 4.12. Default values ranged from 4.00 to 4.50. The calibrated and default PROFACs estimated effort similarly at 62 and 69% within 25%, respectively. The mean calibrated PROFAC had a standard deviation of 1.96, which led to variability in the estimates. Both calibrated and default estimates had mean relative errors in excess of 25%. The variability in the estimates points to analogous records as the first choice in making decisions for PROFAC in new efforts. In the absence of well-defined parameters, use the default values.

PASCAL

The mean calibrated PROFAC, 8.18, was much greater than the mean default PROFAC of 6.42. Default values ranged from 6.00 to 6.50. The mean calibrated PROFAC estimated effort within 25%, 67% of the time versus 17% for default estimates during the most accurate iteration. The six iterations yielded a mean relative error range of 20.89 to 29.11% for calibrated estimates. The recommendation is to use analogous records to achieve the greatest accuracy possible for new estimates; however, the calibrated value of 8.18 should be used in the absence of well-defined new project parameters.

Recommendations for Follow-on Research

The SMC Database is one of the finest automated information sources available in DoD, but it can be improved to enhance future calibration studies. Although each record of the current 2616 contains useful information to some ultimate user, only 70 of those records were found suitable for inclusion in this study. Some of the records were missing

important values crucial to PRICE S calibration such as amount of effort, application mix, and schedule. It would be helpful to update the database records by reviewing the historical archives thoroughly for any missing information to make the database more complete. In instances where crucial parameters are unknown, new research efforts can assist by estimating the parameters, at a minimum, to make more records appropriate for calibration studies.

Perhaps the most significant effort to improve future calibration studies of the SMC Database would involve grouping similar CSCI records together by project and calibrating PROFAC values at the system level. The system PROFAC calibrations can then be used to refine levels of effort reported by the SMC Database at the CSCI level. The visibility into the data gained from this study could lead to less variability in future environmental platform and language calibrations.

Appendix A. Calibration Input Values

Military Ground

Record	PLTFM	Effort	CPLXM	INTEGI	LANG	SIZE	CPLX1	SSR	FCA	APPL	PROFAC
2	1.2	245.0	1.2	0.5	FOR	36200	1.2	187		3.14	3.25
7	1.2	127.0	1.0	0.5	FOR	30160	1.0	1285		6.37	6.74
					ASM	16240					
9	1.2	545.0	1.0	0.7	FOR	117944	1.1	786		7.92	6.73
					ASM	10256					
23	1.2	42.0	1.0	0.5	ATL	5200	1.1	486		8.01	6.28
24	1.2	238.0	1.0	0.5	ATL	18000	1.0	486		8.01	4.98
26	1.2	22.0	1.0	0.5	ATL	4170	1.0	685		8.01	5.79
28	1.2	887.0	1.0	1.0	PL1	107271	1.2	1183		1.98	3.40
					FOR	5646					
30	1.2	160.0	1.2	0.6	PAS	31980	1.1	583		7.48	4.10
					ASM	9020					
48	1.2	344.0	1.0	0.4	FOR	167468	0.9	478		2.16	3.50
50	1.2	656.0	1.0	0.5	FOR	144000	1.0	683	285	7.05	6.35
53	1.2	51.0	1.0	0.5	ASM	22574	1.0	1083		7.46	7.51
54	1.2	134.0	1.0	1.0	PAS	42501	1.2	1083	784	6.19	9.74
					ASM	3199					
55	1.2	6713.0	1.0	1.0	ASM	337432	1.2	1083	784	6.94	2.27
56	1.2	23.0	1.0	0.7	ASM	22482	1.0	1083	784	6.43	7.34
58	1.2	639.0	1.0	0.7	ATL	34672	1.1	1083	384	6.67	4.12
59	1.2	147.0	1.0	0.4	ATL	18578	1.0	1083	384	6.23	6.07
61	1.2	113.0	1.0	0.9	ASM	89417	1.1	1083	484	5.42	9.31
					FOR	6730					
63	1.2	23.0	1.0	0.9	FOR	11753	1.1	1083	484	3.99	7.19
68	1.2	170.0	1.0	0.9	FOR	35140	1.1	1184	685	6.25	6.94
					ASM	18922					
72	1.2	42.0	1.0	0.7	PAS	17914	1.1	885		5.89	8.54
					ASM	7678					
73	1.2	56.0	1.0	0.7	PAS	25645	1.2	1185		5.89	8.09
					ASM	6016					
301	1.2	430.0	1.4	0.4	FOR	90806	0.9	986	1288	4.11	4.88
					ASM	4779					
2497	1.2	80.0	1.4	0.4	ADA	10000	0.9	1090	193	6.75	7.96
2501	1.2	418.0	1.0	0.3	ADA	161700	0.9	891	1192	6.20	6.78
					ASM	3300					
2510	1.2	181.2	1.0	0.4	C	45227	0.9	891	193	5.28	8.42
2519	1.2	1833.0	1.0	0.5	COB	419619	0.9	186	687	2.31	8.22
2520	1.2	3960.0	1.0	0.4	COB	419619	0.8	688	1090	2.31	2.44
2521	1.2	735.0	1.0	0.4	COB	97087	0.8	488	1090	2.31	3.27
2522	1.2	2574.0	1.0	0.4	COB	461426	0.8	588	891	2.31	4.52
2523	1.2	1115.0	1.2	0.4	COB	196365	0.8	1287	389	2.31	6.14
					C	34652					
2524	1.2	1525.0	1.0	0.4	COB	363371	0.8	588	990	2.31	5.52

Military Mobile

Record	PLTFM	Effort	CPLXM	INTEGI	LANG	SIZE	CPLX1	SSR	FCA	APPL	PROFAC
303	1.4	250.0	1.0	0.7	ASM	15000	1.0	1284	1189	8.13	8.59
					PAS	15000					
347	1.4	41.0	1.0	0.7	ADA	2195	1.0	1089		5.64	3.25
					MAC	116					
348	1.4	418.0	1.0	1.0	ADA	16247	1.2	1089		6.94	4.13
					C	1805					
349	1.4	59.0	1.0	0.5	ADA	3104	1.0	1089		8.74	4.55
					MAC	163					
2456	1.4	233.0	1.0	0.7	FOR	81542	1.0	189	991	1.62	7.02
					ADA	7091					
2502	1.4	743.6	1.0	1.0	ADA	53911	1.3	1288	294	8.56	8.46
					ASM	2837					
2507	1.4	759.9	1.0	1.0	ADA	26931	1.3	1288	294	4.79	3.49
					ASM	1417					
2508	1.4	1666.1	1.0	1.0	ADA	59603	1.4	1288	294	8.84	6.30
					ASM	3137					

Missile

Record	PLTFM	Effort	CPLXM	INTEGI	LANG	SIZE	CPLX1	SSR	FCA	APPL	PROFAC
15	1.8	276.0	1.0	0.7	ASM	11136	1.2	587		8.46	3.08
16	1.8	96.0	1.0	0.7	ASM	13207	1.1	587		8.46	5.86
27	1.8	1460.0	1.0	0.7	JOV	18933	1.2	1283		5.67	5.37
36	1.8	506.0	1.2	1.0	ASM	13658	1.3	281		5.61	4.99

Unmanned Space

Record	PLTFM	Effort	CPLXM	INTEGI	LANG	SIZE	CPLX1	SSR	FCA	APPL	PROFAC
29	2.0	52.0	1.0	0.5	ADA	1900	1.1	688		10.95	4.21
					ASM	100					
38	2.0	64.0	1.0	1.0	ASM	6000	1.4	586		2.85	2.66
74	2.0	80.0	1.0	0.7	JOV	11700	1.1	1082	385	3.69	3.03
75	2.0	912.0	1.0	0.7	JOV	116800	1.1	581	585	7.49	5.56
76	2.0	115.0	1.0	0.5	JOV	14000	1.0	381	984	4.36	3.29
77	2.0	523.0	1.0	0.7	JOV	56200	1.3	581	285	3.94	2.48
78	2.0	478.0	1.0	1.0	JOV	48300	1.3	681	385	6.16	4.46
79	2.0	432.0	1.0	1.0	JOV	50300	0.9	581	185	5.67	3.67
80	2.0	296.0	1.0	0.7	FOR	31253	1.1	981	785	2.12	3.02
					JOV	38197					
81	2.0	164.0	1.0	0.7	JOV	22900	1.1	381	684	6.40	4.70
82	2.0	140.0	1.0	0.7	JOV	16300	1.1	381	1085	3.65	2.86
83	2.0	57.0	1.0	0.7	JOV	6800	1.1	381	885	3.74	3.35
305	2.0	145.0	1.2	0.3	ADA	6186	0.8	990		8.59	8.66
					C	14434					
306	2.0	90.0	1.0	0.7	C	9500	1.1	887		7.63	8.41

ADA

Record	PLTFM	Effort	CPLXM	INTEGI	LANG	SIZE	CPLX1	SSR	FCA	APPL	PROFAC
29	2.0	52.0	1.0	0.5	ADA	1900	1.1	688		10.95	4.25
					ASM	100					3.80
347	1.4	41.0	1.0	0.7	ADA	2195	1.0	1089		5.64	3.30
					MAC	116					2.10
348	1.4	418.0	1.0	1.0	ADA	16247	1.2	1089		6.94	4.13
					C	1805					4.13
349	1.4	59.0	1.0	0.5	ADA	3104	1.0	1089		8.74	4.60
					MAC	163					3.00
2497	1.2	80.0	1.4	0.4	ADA	10000	0.9	1090	193	6.75	7.96
2501	1.2	418.0	1.0	0.3	ADA	161700	0.9	891	1192	6.20	6.78
					ASM	3300					6.40
2502	1.4	743.6	1.0	1.0	ADA	53911	1.3	1288	294	8.56	8.46
					ASM	2837					8.35
2507	1.4	759.9	1.0	1.0	ADA	26931	1.3	1288	294	4.79	3.50
					ASM	1417					3.30
2508	1.4	1666.1	1.0	1.0	ADA	59603	1.4	1288	294	8.84	6.30
					ASM	3137					6.20
2512	1.8	259.0	1.0	0.7	ADA	32495	1.1	691	193	7.22	6.25
					ASM	663					6.10

ASSEMBLY

Record	PLTFM	Effort	CPLXM	INTEGI	LANG	SIZE	CPLX1	SSR	FCA	APPL	PROFAC
15	1.8	276.0	1.0	0.7	ASM	11136	1.2	587		8.46	3.08
16	1.8	96.0	1.0	0.7	ASM	13207	1.1	587		8.46	5.86
36	1.8	506.0	1.2	1.0	ASM	13658	1.3	281		5.61	4.99
38	2.0	64.0	1.0	1.0	ASM	6000	1.4	586		2.85	2.66
53	1.2	51.0	1.0	0.5	ASM	22574	1.0	1083		7.46	7.51
55	1.2	6713.0	1.0	1.0	ASM	337432	1.2	1083	784	6.94	2.27
56	1.2	23.0	1.0	0.7	ASM	22482	1.0	1083	784	6.43	7.34
61	1.2	113.0	1.0	0.9	ASM	89417	1.1	1083	484	5.42	9.20
					FOR	6730					9.60
302	1.8	400.0	1.0	0.6	ASM	26000	1.1	488	1189	7.19	8.55
					JOV	26000					8.45
303	1.4	250.0	1.0	0.7	ASM	15000	1.0	1284	1189	8.13	7.70
					PAS	15000					8.90

C

Record	PLTFM	Effort	CPLXM	INTEGI	LANG	SIZE	CPLX1	SSR	FCA	APPL	PROFAC
305	2.0	145.0	1.2	0.3	ADA	6186	0.8	990		8.59	8.66
					C	14434					8.66
306	2.0	90.0	1.0	0.7	C	9500	1.1	887		7.63	8.41
307	1.2	52.0	1.2	0.3	C	25000	0.8	586	787	5.21	8.87
2510	1.2	181.2	1.0	0.4	C	45227	0.9	891	193	5.28	8.42

COBOL

Record	PLTFM	Effort	CPLXM	INTEGI	LANG	SIZE	CPLX1	SSR	FCA	APPL	PROFAC
2519	1.2	1833.0	1.0	0.5	COB	419619	0.9	186	687	2.31	8.22
2520	1.2	3960.0	1.0	0.4	COB	419619	0.8	688	1090	2.31	2.44
2521	1.2	735.0	1.0	0.4	COB	97087	0.8	488	1090	2.31	3.27
2522	1.2	2574.0	1.0	0.4	COB	461426	0.8	588	891	2.31	4.52
2523	1.2	1115.0	1.2	0.4	COB	196365	0.8	1287	389	2.31	6.20
					C	34652					5.75
2524	1.2	1525.0	1.0	0.4	COB	363371	0.8	588	990	2.31	5.52

FORTRAN

Record	PLTFM	Effort	CPLXM	INTEGI	LANG	SIZE	CPLX1	SSR	FCA	APPL	PROFAC
2	1.2	245.0	1.2	0.5	FOR	36200	1.2	187		3.14	3.25
7	1.2	127.0	1.0	0.5	FOR	30160	1.0	1285		6.37	6.80
					ASM	16240					6.50
9	1.2	545.0	1.0	0.7	FOR	117944	1.1	786		7.92	6.75
					ASM	10256					6.00
48	1.2	344.0	1.0	0.4	FOR	167468	0.9	478		2.16	3.50
50	1.2	656.0	1.0	0.5	FOR	144000	1.0	683	285	7.05	6.35
63	1.2	23.0	1.0	0.9	FOR	11753	1.1	1083	484	3.99	7.19
68	1.2	170.0	1.0	0.9	FOR	35140	1.1	1184	685	6.25	7.00
					ASM	18922					6.55
80	2.0	296.0	1.0	0.7	FOR	31253	1.1	981	785	2.12	3.40
					JOV	38197					2.70
301	1.2	430.0	1.4	0.4	FOR	90806	0.9	986	1288	4.11	4.90
					ASM	4779					3.90
2456	1.4	233.0	1.0	0.7	FOR	81542	1.0	189	991	1.62	7.05
					ADA	7091					6.70

JOVIAL

Record	PLTFM	Effort	CPLXM	INTEGI	LANG	SIZE	CPLX1	SSR	FCA	APPL	PROFAC
14	1.8	490.0	1.0	0.7	JOV	22148	1.2	784		4.01	3.44
27	1.8	1460.0	1.0	0.7	JOV	18933	1.2	1283		5.67	5.37
74	2.0	80.0	1.0	0.7	JOV	11700	1.1	1082	385	3.69	3.03
75	2.0	912.0	1.0	0.7	JOV	116800	1.1	581	585	7.49	5.56
76	2.0	115.0	1.0	0.5	JOV	14000	1.0	381	984	4.36	3.29
77	2.0	523.0	1.0	0.7	JOV	56200	1.3	581	285	3.94	2.48
78	2.0	478.0	1.0	1.0	JOV	48300	1.3	681	385	6.16	4.46
79	2.0	432.0	1.0	1.0	JOV	50300	0.9	581	185	5.67	3.67
80	2.0	296.0	1.0	0.7	FOR	31253	1.1	981	785	2.12	3.40
					JOV	38197					2.70
81	2.0	164.0	1.0	0.7	JOV	22900	1.1	381	684	6.40	4.70
82	2.0	140.0	1.0	0.7	JOV	16300	1.1	381	1085	3.65	2.86
83	2.0	57.0	1.0	0.7	JOV	6800	1.1	381	885	3.74	3.35
302	1.8	400.0	1.0	0.6	ASM	26000	1.1	488	1189	7.19	8.55
					JOV	26000					8.45

PASCAL

Record	PLTFM	Effort	CPLXM	INTEGI	LANG	SIZE	CPLX1	SSR	FCA	APPL	PROFAC
30	1.2	160.0	1.2	0.6	PAS	31980	1.1	583		7.48	4.10
					ASM	9020					4.05
54	1.2	134.0	1.0	1.0	PAS	42501	1.2	1083	784	6.19	9.80
					ASM	3199					8.20
70	1.2	40.0	1.0	0.7	PAS	7008	1.2	185	1085	5.11	5.30
					ASM	292					3.70
72	1.2	42.0	1.0	0.7	PAS	17914	1.1	885		5.89	8.70
					ASM	7678					7.20
73	1.2	56.0	1.0	0.7	PAS	25645	1.2	1185		5.89	8.20
					ASM	6016					6.80
303	1.4	250.0	1.0	0.7	ASM	15000	1.0	1284	1189	8.13	7.70
					PAS	15000					8.90

Appendix B. Wilcoxon Signed-Rank Tests

Military Ground

Record #	Actual Effort	Cal Effort	Diff	Rank	Def Effort	Diff	Rank
2	245.0	125.10	-119.90	22.0	128.10	-116.90	22.0
7	127.0	140.00	13.00	8.0	142.90	15.90	7.0
9	545.0	600.30	55.30	16.0	613.20	68.20	17.0
23	42.0	43.00	1.00	1.0	43.90	1.90	3.0
24	238.0	195.10	-42.90	14.0	198.70	-39.30	13.0
26	22.0	20.70	-1.30	2.5	21.10	-0.90	1.0
28	887.0	479.00	-408.00	27.0	488.70	-398.30	27.0
30	160.0	103.60	-56.40	17.0	105.80	-54.20	15.0
48	344.0	178.40	-165.60	24.0	182.50	-161.50	24.0
50	656.0	682.50	26.50	12.0	697.20	41.20	14.0
53	51.0	62.50	11.50	7.0	63.70	12.70	6.0
54	134.0	235.90	101.90	21.0	241.40	107.40	21.0
55	6713.0	2390.80	-4322.20	31.0	2439.80	-4273.20	31.0
56	23.0	27.90	4.90	5.0	28.50	5.50	4.0
58	639.0	419.70	-219.30	25.0	428.40	-210.60	25.0
59	147.0	145.70	-1.30	2.5	148.70	1.70	2.0
61	113.0	186.90	73.90	19.0	191.30	78.30	19.0
63	23.0	28.00	5.00	6.0	28.70	5.70	5.0
68	170.0	197.20	27.20	13.0	201.80	31.80	12.0
72	42.0	60.50	18.50	9.0	61.70	19.70	8.0
73	56.0	75.20	19.20	10.0	76.80	20.80	9.0
301	430.0	339.30	-90.70	20.0	346.10	-83.90	20.0
2497	80.0	103.30	23.30	11.0	105.10	25.10	10.0
2501	418.0	467.70	49.70	15.0	477.40	59.40	16.0
2510	181.2	247.30	66.10	18.0	251.90	70.70	18.0
2519	1833.0	2547.90	714.90	29.0	2602.30	769.30	29.0
2520	3960.0	1575.70	-2384.30	30.0	1611.30	-2348.70	30.0
2521	735.0	388.90	-346.10	26.0	397.90	-337.10	26.0
2522	2574.0	1892.50	-681.50	28.0	1935.40	-638.60	28.0
2523	1115.0	1118.60	3.60	4.0	1143.30	28.30	11.0
2524	1525.0	1371.10	-153.90	23.0	1402.10	-122.90	23.0

	Cal Effort	Def Effort	n	T(a=.10)	T(a=.05)
Pos Dif	**204**	**211**	31	163	148
Neg Dif	292	285			
a=.10	no bias	no bias			
a=.05	no bias	no bias			

Military Mobile

Record #	Actual Effort	Cal Effort	Diff	Rank	Def Effort	Diff	Rank
303	250.0	370.30	120.30	5.0	384.50	134.50	5.0
347	41.0	23.30	-17.70	2.0	24.20	-16.80	2.0
348	418.0	304.30	-113.70	4.0	316.20	-101.80	4.0
349	59.0	47.20	-11.80	1.0	49.10	-9.90	1.0
2456	233.0	289.70	56.70	3.0	302.10	69.10	3.0
2502	743.6	1089.60	346.00	8.0	1131.90	388.30	8.0
2507	759.9	472.10	-287.80	7.0	490.50	-269.40	7.0
2508	1666.1	1822.50	156.40	6.0	1890.50	224.40	6.0

	Cal Effort	Def Effort	n	T(a=.10)	T(a=.05)
Pos Dif	22	22	8	6	4
Neg Dif	14	14			
a=.10	no bias	no bias			
a=.05	no bias	no bias			

Unmanned Space

Record #	Actual Effort	Cal Effort	Diff	Rank	Def Effort	Diff	Rank
29	52.0	50.60	-1.40	1.0	48.70	-3.30	1.0
38	64.0	39.50	-24.50	6.0	37.80	-26.20	6.0
74	80.0	56.70	-23.30	5.0	54.60	-25.40	5.0
75	912.0	1212.30	300.30	14.0	1160.60	248.60	14.0
76	115.0	88.30	-26.70	7.0	85.10	-29.90	7.0
77	523.0	301.70	-221.30	13.0	290.20	-232.80	13.0
78	478.0	495.20	17.20	4.0	473.70	-4.30	2.0
79	432.0	367.10	-64.90	9.0	353.00	-79.00	10.0
80	296.0	206.90	-89.10	11.0	199.00	-97.00	11.0
81	164.0	178.10	14.10	3.0	171.70	7.70	3.0
82	140.0	94.10	-45.90	8.0	90.60	-49.40	8.0
83	57.0	44.50	-12.50	2.0	42.80	-14.20	4.0
305	145.0	287.70	142.70	12.0	277.10	132.10	12.0
306	90.0	172.80	82.80	10.0	166.40	76.40	9.0

	Cal Effort	Def Effort	n	T(a=.10)	T(a=.05)
Pos Dif	43	38	14	26	21
Neg Dif	62	67			
a=.10	no bias	no bias			
a=.05	no bias	no bias			

ADA

Record #	Actual Effort	Cal Effort	Diff	Rank	Def Effort	Diff	Rank
29	52.0	40.60	-11.40	2.0	44.60	-7.40	1.0
347	41.0	24.50	-16.50	3.0	22.50	-18.50	4.0
348	418.0	323.60	-94.40	6.0	299.50	-118.50	8.0
349	59.0	49.30	-9.70	1.0	45.50	-13.50	2.0
2497	80.0	113.50	33.50	4.0	97.00	17.00	3.0
2501	418.0	520.50	102.50	7.0	437.40	19.40	5.0
2502	743.6	1118.60	375.00	10.0	1032.20	288.60	9.0
2507	759.9	493.50	-266.40	9.0	454.70	-305.20	10.0
2508	1666.1	1882.80	216.70	8.0	1741.70	75.60	7.0
2512	259.0	299.20	40.20	5.0	299.80	40.80	6.0

	Cal Effort	Def Effort	n	T(a=.10)	T(a=.05)
Pos Dif	34	30	10	11	8
Neg Dif	**21**	**25**			
a=.10	no bias	no bias			
a=.05	no bias	no bias			

ASSEMBLY

Record #	Actual Effort	Cal Effort	Diff	Rank	Def Effort	Diff	Rank
15	276.0	145.10	-130.90	9.0	193.80	-82.20	9.0
16	96.0	92.60	-3.40	1.0	124.20	28.20	5.0
36	506.0	414.90	-91.10	8.0	555.60	49.60	6.0
38	64.0	27.90	-36.10	5.0	42.60	-21.40	3.0
53	51.0	62.70	11.70	3.0	69.60	18.60	2.0
55	6713.0	2398.80	-4314.20	10.0	2685.80	-4027.20	10.0
56	23.0	28.00	5.00	2.0	31.40	8.40	1.0
61	113.0	167.10	54.10	7.0	182.30	69.30	8.0
302	400.0	437.60	37.60	6.0	469.10	69.10	7.0
303	250.0	262.80	12.80	4.0	272.30	22.30	4.0

	Cal Effort	Def Effort	n	T(a=.10)	T(a=.05)
Pos Dif	**22**	33	10	11	8
Neg Dif	33	**22**			
a=.10	no bias	no bias			
a=.05	no bias	no bias			

FORTRAN

Record #	Actual Effort	Cal Effort	Diff	Rank	Def Effort	Diff	Rank
2	245.0	137.50	-107.50	8.0	117.60	-127.40	9.0
7	127.0	150.10	23.10	2.0	131.70	4.70	2.0
9	545.0	657.80	112.80	9.0	565.80	20.80	5.0
48	344.0	197.80	-146.20	10.0	166.60	-177.40	10.0
50	656.0	753.20	97.20	7.0	639.60	-16.40	4.0
63	23.0	31.20	8.20	1.0	26.10	3.10	1.0
68	170.0	212.90	42.90	3.0	183.70	13.70	3.0
80	296.0	219.00	-77.00	6.0	236.80	-59.20	7.0
301	430.0	373.60	-56.40	4.0	321.90	-108.10	8.0
2456	233.0	291.10	58.10	5.0	273.00	40.00	6.0

	Cal Effort	Def Effort	n	T(a=.10)	T(a=.05)
Pos Dif	27	17	10	11	8
Neg Dif	28	38			
a=.10	no bias	no bias			
a=.05	no bias	no bias			

JOVIAL

Record #	Actual Effort	Cal Effort	Diff	Rank	Def Effort	Diff	Rank
14	490.0	393.30	-96.70	9.0	379.60	-110.40	9.0
27	1460.0	1792.70	332.70	13.0	1731.90	271.90	12.0
74	80.0	56.50	-23.50	4.0	61.30	-18.70	2.0
75	912.0	1209.30	297.30	12.0	1315.50	403.50	13.0
76	115.0	88.10	-26.90	5.0	94.70	-20.30	3.0
77	523.0	301.00	-222.00	10.0	324.70	-198.30	10.0
78	478.0	494.00	16.00	3.0	538.50	60.50	8.0
79	432.0	366.30	-65.70	7.0	395.20	-36.80	5.0
80	296.0	230.20	-65.80	8.0	240.60	-55.40	7.0
81	164.0	177.70	13.70	2.0	190.90	26.90	4.0
82	140.0	93.90	-46.10	6.0	101.00	-39.00	6.0
83	57.0	44.40	-12.60	1.0	47.80	-9.20	1.0
302	400.0	661.80	261.80	11.0	644.10	244.10	11.0

	Cal Effort	Def Effort	n	T(a=.10)	T(a=.05)
Pos Dif	41	48	13	21	17
Neg Dif	50	43			
a=.10	no bias	no bias			
a=.05	no bias	no bias			

Bibliography

1. Blackmar, John C. Avionics Software Cost/Schedule Model Optimization. ASD Reserve Project 81-146-PIT. March 1982.

2. Boehm, Barry W. Software Engineering Economics. Englewood Cliffs: Prentice-Hall, 1981.

3. Brooks, Frederick P., Jr. The Mythical Man-Month. Reading: Addison-Wesley Publishing Company, 1975.

4. Coggins, George A. and Roy C. Russell. Software Cost Estimating Models: A Comparative Study of What the Models Estimate. MS thesis, AFIT/GCA/LAS/93S-4. School of Logistics and Acquisition Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1993 (AD-A275989).

5. Conte, Samuel D., and others. Software Engineering Metrics and Models. Menlo Park: Benjamin/Cummings, 1986.

6. Ferens, Daniel V. Class handouts, Cost 677 Note-Taking Devices. School of Logistics and Acquisition Management, Air Force Institute of Technology, Wright-Patterson AFB OH, Fall 1994.

7. Ferens, Daniel V. Class handout, Cost 677 Readings. School of Logistics and Acquisition Management, Air Force Institute of Technology, Wright-Patterson AFB OH, Fall 1994.

8. GE PRICE Systems. PRICE S Calibration: Technical Bulletin No. 14.

9. GE PRICE Systems. The Central Equations of the PRICE Software Cost Model.

10. James, Thomas G. Jr., and Daniel V. Ferens. Application of the RCA PRICE-S Software Cost Estimation Model to Air Force Avionics Laboratory Programs. Avionics Systems Engineering Branch, Wright-Patterson AFB OH, October 1979 (AF-TR-79-1164).

11. Londeix, Bernard. Cost Estimating for Software Development. New York: Addison-Wesley Publishing Company, 1987.

12. Management Consulting and Research. Air Force Cost Analysis Agency Software Model Content Study. Oxnard: Management Consulting and Research, 1994.

13. Management Consulting and Research. Application Oriented Software Data Collection Software Model Calibration Report. Oxnard: Management Consulting and Research, 1991.

14. Management Consulting and Research. Space and Missile Systems Center Software Database User's Manual. Oxnard: Management Consulting and Research, 1993.

15. Mendenhall, William and others. Mathematical Statistics with Applications. Belmont: Duxbury Press, 1990.

16. Ourada, Gerald L. Software Cost Estimating Models: A Calibration, Validation, and Comparison. MS thesis, AFIT/GSS/LSY/91D-11. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, December 1991 (AD-A246677).

17. PRICE S. Release 2.0. Computer software. PRICE Systems, Moorestown NJ, 1992, 1993.

18. PRICE Systems. Calibration of the PRICE S Model.

19. PRICE Systems. PRICE S Reference Manual. Moorestown: PRICE Systems. 1987.

20. PRICE Systems. PRICE S Reference Manual. Moorestown: PRICE Systems. 1989.

21. PRICE Systems. PRICE S Reference Manual. 3rd ed. Moorestown: PRICE Systems. 1993.

22. RCA Aerospace and Defense PRICE Systems. "Interim Report Number Five for Calibration of the RCA PRICE Models for the Deputy for Reconnaissance and Warfare." Aeronautical Systems Division, Wright-Patterson AFB OH, February 29, 1988 (F33600-84-F4504).

23. RCA Aerospace and Defense PRICE Systems. "Phase II Technical Report for Calibration of the PRICE Models for the Deputy for Reconnaissance and Warfare." Aeronautical Systems Division, Wright-Patterson AFB OH, October 31, 1988 (F33600-84-F4504).

24. Smith, Palmer W. "The Challenges of Costing Very Large Software Development Programs." NES/ICA 2nd Annual Cost Symposium, April 18-19, 1989.

25. Space and Missile Systems Center. Software Database Data Collection Dictionary. 1993.

26. Symons, Charles R. Software Sizing and Estimating: Mk II FPA. New York: John Wiley & Sons, 1991.

27. Thibodeau, Robert. An Evaluation of Software Cost Estimating Models. Rome Air Development Center, Air Force Systems Command, Griffiss AFB NY, September 1991 (AD-A104226).

28. Wellman, Frank. Software Costing. New York: Prentice-Hall, 1992.

Vita

Captain James C. Galonsky was born on 3 June 1965 at Andrews Air Force Base, Maryland. He grew up in Virginia Beach, VA and graduated with Honors from Frank W. Cox High School in 1983. As a junior at Virginia Polytechnic Institute and State University, he enrolled in the Virginia Tech Corps of Cadets (VTCC) and entered the two-year Air Force ROTC program. He received his commission on 20 May 1988 and earned a Bachelor of Science degree in Finance. His first assignment was to Vandenberg AFB for training on the Improved Launch Control System (ILCS) of the Minuteman II weapon system. His missile operations tour was spent at Whiteman AFB, MO where he served as the Senior Flight Commander of the 508th Missile Squadron and the Commander of the Training and Evaluation Flight. Upon completion of his controlled tour, he attended Squadron Officer School in residence. In April 1994, he entered the School of Logistics and Acquisition Management, Air Force Institute of Technology.

Permanent Address: 2305 Inlynnview Rd
Virginia Beach, VA 23454

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (*Leave blank*)

2. REPORT DATE
September 1995

3. REPORT TYPE AND DATES COVERED
Master's Thesis

4. TITLE AND SUBTITLE
CALIBRATION OF THE PRICE S SOFTWARE COST MODEL

5. FUNDING NUMBERS

6. AUTHOR(S)
James C. Galonsky, Captain USAF

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology,
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/GCA/LAS/95S-1

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
HQ SMC/FMC
2430 E El Segundo Blvd #2010
El Segundo CA 90245-4687

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)

As more Department of Defense resources are being allocated toward software development, the necessity to accurately plan for software costs has become critical. Obtaining reliable estimates from software cost models, like PRICE S, can be a problem when input parameters are not precisely defined or calibrated. This research effort centered on refining Productivity Factor (PROFAC) values for defense industry applications. The Space and Missile Systems Center Database was used to calibrate PROFAC values for eleven stratified data sets: military ground, military mobile, missile, unmanned space, Ada, Assembly, C, COBOL, FORTRAN, JOVIAL, and PASCAL. The accuracy of the calibrations was determined through comparisons of calibrated and default generated estimates versus actuals. The results were surprising in that, in most instances, the calibrated PROFAC values estimated actual cost well, but not overwhelmingly better than the default PROFAC values. The main contributing factor to this phenomena was variability within the stratified data sets. The results were encouraging, however, in that the results from seven of the eleven stratified data sets suggested either a new refinement in PROFAC values based upon the calibration or the recommendation to use PROFAC values from analogous calibrated records for estimating future efforts.

14. SUBJECT TERMS
Cost Estimates, Cost Models,
Calibration, Software Engineering

15. NUMBER OF PAGES
93

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL